**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/15, C07K 19/00, 14/81, C12N 9/54, 1/21, C11D 3/386 — A1

(11) International Publication Number: WO 97/15670

(43) International Publication Date: 1 May 1997 (01.05.97)

(21) International Application Number: PCT/US96/17153

(22) International Filing Date: 25 October 1996 (25.10.96)

(30) Priority Data:
08/548,186 25 October 1995 (25.10.95) US

(71) Applicant: ARRIS PHARMACEUTICAL CORPORATION [US/US]; Suite 3, 385 Oyster Point Boulevard, South San Francisco, CA 94080 (US).

(72) Inventors: HARTMAN, Christa; 4607 Walnut Street, Oakland, CA 94619 (US). ESTELL, David, A.; 248 Woodbridge Circle, San Mateo, CA 94403 (US). SCHMIDT, Brian; 467 Filbert Street, Half Moon Bay, CA 94019 (US). MCGRATH, Mary; 1540 Balboa Way, Burlingame, CA 94010 (US). FISHER, Joseph; 1035 D. South Van Ness, San Francisco, CA 94110 (US).

(74) Agents: BREZNER, David, J. et al.; Flehr, Hohbach, Test, Albritton & Herbert, Suite 3400, 4 Embarcadero Center, San Francisco, CA 94111-4187 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

**Published**
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: NOVEL SUBTILISIN INHIBITORS

(57) Abstract

The invention relates to novel subtilisin inhibitors comprising a human standard mechanism inhibitor that has the reactive site of inhibition replaced with the reactive site from a known subtilisin inhibitor. The invention provides nucleic acids, proteins, and methods for inhibiting subtilisins, and detergents containing modified subtilisin inhibitors.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| **AM** | Armenia | **GB** | United Kingdom | **MW** | Malawi |
| **AT** | Austria | **GE** | Georgia | **MX** | Mexico |
| **AU** | Australia | **GN** | Guinea | **NE** | Niger |
| **BB** | Barbados | **GR** | Greece | **NL** | Netherlands |
| **BE** | Belgium | **HU** | Hungary | **NO** | Norway |
| **BF** | Burkina Faso | **IE** | Ireland | **NZ** | New Zealand |
| **BG** | Bulgaria | **IT** | Italy | **PL** | Poland |
| **BJ** | Benin | **JP** | Japan | **PT** | Portugal |
| **BR** | Brazil | **KE** | Kenya | **RO** | Romania |
| **BY** | Belarus | **KG** | Kyrgystan | **RU** | Russian Federation |
| **CA** | Canada | **KP** | Democratic People's Republic of Korea | **SD** | Sudan |
| **CF** | Central African Republic | | | **SE** | Sweden |
| **CG** | Congo | **KR** | Republic of Korea | **SG** | Singapore |
| **CH** | Switzerland | **KZ** | Kazakhstan | **SI** | Slovenia |
| **CI** | Côte d'Ivoire | **LI** | Liechtenstein | **SK** | Slovakia |
| **CM** | Cameroon | **LK** | Sri Lanka | **SN** | Senegal |
| **CN** | China | **LR** | Liberia | **SZ** | Swaziland |
| **CS** | Czechoslovakia | **LT** | Lithuania | **TD** | Chad |
| **CZ** | Czech Republic | **LU** | Luxembourg | **TG** | Togo |
| **DE** | Germany | **LV** | Latvia | **TJ** | Tajikistan |
| **DK** | Denmark | **MC** | Monaco | **TT** | Trinidad and Tobago |
| **EE** | Estonia | **MD** | Republic of Moldova | **UA** | Ukraine |
| **ES** | Spain | **MG** | Madagascar | **UG** | Uganda |
| **FI** | Finland | **ML** | Mali | **US** | United States of America |
| **FR** | France | **MN** | Mongolia | **UZ** | Uzbekistan |
| **GA** | Gabon | **MR** | Mauritania | **VN** | Viet Nam |

# NOVEL SUBTILISIN INHIBITORS

## FIELD OF THE INVENTION

The invention relates to novel subtilisin inhibitors comprising a human standard mechanism inhibitor that has the reactive site of inhibition replaced with the reactive site from a known subtilisin inhibitor. The invention provides nucleic acids, proteins, and methods for inhibiting subtilisins, and detergents containing modified subtilisin inhibitors.

## BACKGROUND OF THE INVENTION

Several serine proteases have use in commercial processes. For example, subtilisin, a serine protease from Bacillus species, is used extensively in laundry detergent formulations. Detergent formulations containing enzymes, and particularly subtilisin, may exhibit enzyme stability problems upon storage, especially liquid detergents which can contain high levels of surfactants and water. An enzyme may hydrolyze in water and may degrade itself or other enzymes, resulting in the addition of extra enzyme or other costly additives to retard degradation.

Thus a variety of attempts have been made to stabilize detergent formulations to avoid the loss of enzymatic activity. See for example U.S. Patent Nos.

5,039,446, and 5,178,789, and cited documents, and European Patent Nos. 0 199 405 B1 and 0 549 675 B1.

In addition, the allergenicity of the enzyme is also problematic for workers involved in producing the enzyme and the detergents containing them. In the case of subtilisin, particular allergenicity is associated with the active enzyme, which may cause proteolytic degradation of other proteins, including host proteins, leading to an immune response. Thus many techniques to limit exposure of workers to the active enzyme have been explored, including granularization, immobilization and inhibition of the enzyme, in an attempt to avoid the release of proteinaceous material during normal handling and storage.

However, the added inhibitor may itself be allergenic, such as the case with the ovomucoid third domain inhibitors, which are potent inhibitors of subtilisin.

Therefore, it is desirable to make a subtilisin protease inhibitor which does not itself cause a significant allergic reaction. WO 92/10755 describes methods for selecting the immunological epitopes of a protein for the purpose of changing the sequence and reducing the immunological response.

Accordingly, it is an object of the invention to provide subtilisin inhibitors which may be used to inhibit subtilisin to prevent or decrease an immune response to the inhibitor and the enzyme-inhibitor complex.

## SUMMARY OF THE INVENTION

In accordance with the foregoing objects, it is an aspect of the invention to provide modified subtilisin inhibitors comprising a modified human standard mechanism inhibitor. At least one amino acid of the reactive site of a human standard mechanism inhibitor is substituted with a different amino acid, such

that the $K_I$ of the modified subtilisin inhibitor against at least one subtilisin is decreased by at least a factor of 100 relative to the unmodified human standard mechanism inhibitor.

Also provided are modified subtilisin inhibitors wherein at least five amino acids of the reactive site of the human standard mechanism inhibitor are substituted with different amino acids.

Also provided are modified subtilisin inhibitors wherein the reactive site of the human standard mechanism inhibitor is replaced by the reactive site of a known subtilisin inhibitor.

Also provided are methods of inhibiting a target subtilisin comprising reversibly binding to a target subtilisin a modified subtilisin inhibitor. The modified subtilisin inhibitor comprises a human standard mechanism inhibitor wherein the reactive site of the human standard mechanism inhibitor is replaced by the reactive site of a known subtilisin inhibitor. The $K_I$ of the modified subtilisin inhibitor against at least one subtilisin is decreased by at least a factor of 100 relative to the unmodified human standard mechanism inhibitor.

Further provided are target subtilisins inhibited by a modified subtilisin inhibitor of the invention.

Additionally provided are nucleic acids encoding a modified human standard mechanism inhibitor, wherein at least one amino acid of the reactive site of said human standard mechanism inhibitor is substituted with a different amino acid, such that the $K_I$ of the modified subtilisin inhibitor against at least one subtilisin is decreased by at least a factor of 100 relative to the unmodified human standard mechanism inhibitor. Also provided are nucleic acids wherein the

reactive site of the human standard mechanism inhibitor is replaced by the reactive site of a known subtilisin inhibitor.

Further provided are expression vectors and host cells comprising the nucleic acids of the invention.

Additionally provided are detergent compositions. The compositions comprise a modified subtilisin inhibitor of the invention and a detergent containing a target subtilisin.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts the pBN3 plasmid with the restriction sites shown, as described in Example 1.

Figure 2 depicts the pKKphoA.PSTI.SSI.M2 plasmid as described in Example 2.

Figure 3 depicts the standard mechansims inhibitors hPSTI, and acrosin inhibitor (depicted in the Figure as HUSI-II). Human fooistatin (FOLL), human agrin domain 7 (hAGR7) and human hevin (HEVIN) are also depicted; they all have sequence homology to hPSTI. They could also be used as backbones (scaffolds) for the modified protease inhibitors of the invention, although currently they have not been shown to inhibit proteases without additional modifications. The P10 to P10' sites and the P1-P1' junctions are depicted (known for hPSTI and acrosin inhibitor; putative identification based on homology for the others). It should be noted for all of the sequences depicted in the Figures that the P10 to P10' region is determined by homology, and thus the individual inhibitors may have a few more or a few less residues than the 20

residues of the P10 to P10' site generally depicted. That is, the P10 to P10' site is determined by homology to the hPSTI site depicted on the top line.

Figure 4 depicts the MPI-1, MPI-2 (domains 1 and 2 of mucous protease inhibitor), HE4-1, HE4-2 (domains 1 and 2 of human epididymis protease inhibitor) and SKALP inhibitors, with the P10 to P10' and P1-P1' junctions depicted. As above, the P10 to P10' site may have a few more or a few less residues.

Figure 5 depicts the ovomucoid third domains of 106 avian species, with the P10 to P10' and P1-P1' junctions depicted. These sequences are originally reported in Laskowski et al., Biochem. 26:202-221 (1987), hereby expressly incorporated by reference (additional sequences were submitted to the Protein Sequence Database; see 1990).

Figure 6 depicts the reactive sites of the SSI inhibitors, with the P10 to P10' and P1-P1' sites shown. As above, the P10 to P10' site is determined on the basis of homology to the top SSI sequence, and thus the actual P10 to P10' site for some of the other inhibitors may have a few more or a few less residues in the lineup.

Figures 7A and 7B depict the sequences of the red sea turtle inhibitor and rat agrin domain 5, with the P10 to P10' sites and the P1-P1' junctions shown based on homology to MPI-2 and hPSTI, respectively.

Figure 8 depicts the sequence of the hPSTI.M2 gene for baculoviral expression, with a BamHI site and a EcoRI site shown.

Figures 9A and 9B depict the sequence of the hPSTI.SSI inhibitor. Figure 9A depicts the hPSTI.SSI.M2 construction, for expression in *E. coli* using the

phoA promoter, with restriction sites shown. The signal sequence is underlined, the reactive site from SSI is shown in bold, and the M2 tag is in italics. The first alanine after the signal sequence is from the cloning vector, and is not necessary. Figure 9B depicts the amino acid sequence of the hPSTI.SSI modified subtilisin inhibitor without the M2 tag or the N-terminal alanine.

Figures 10A and 10B depict the sequence of the hPSTI.OMTKY3 modified subtilisin inhibitor. Figure 10A the sequence of the hPSTI.OMTKY3.M2 construction, with restriction sites shown. The signal sequence is underlined, the reactive site from SSI is shown in bold, and the M2 tag is in italics. Figure 10B depicts the amino acid sequence of the hPSTI.OMTKY3 modified subtilisin inhibitor without the M2 tag or the N-terminal alanine.

Figures 11A and 11B depict the sequence of the hPSTI.SSImutant modified subtilisin inhibitor. Figure 11A depicts the sequence of the hPSTI.SSImutant.M2 construction, with restriction sites shown. The signal sequence is underlined, the reactive site from SSI is shown in bold, and the M2 tag is in italics. Figure 10B depicts the amino acid sequence of the hPSTI.SSImutant modified subtilisin inhibitor without the M2 tag or the N-terminal alanine.

Figure 12 depicts the sequences inserted into pBN3 to make pBNppt, as described in the Examples. These are the signal sequence and prosequence of a subtilisin for use in cloning.

Figure 13 depicts the sequence of the silkworm sequence, with a putative P1-P1' site identified.

Figure 14 depicts the rat agrin domain 5 sequence used to express in baculovirus.

## DETAILED DESCRIPTION OF THE INVENTION

The present invention provides novel modified serine protease inhibitors, and in particular, novel modified subtilisin inhibitors. Generally, known human standard mechanism inhibitors do not exhibit suitable kinetic characteristics to allow their use in commercial processes involving subtilisin, although they do not generate undesirable immune responses. Similarly, known subtilisin inhibitors may have appropriate kinetic properties but are too allergenic or otherwise unsuitable for commercial use with humans. Accordingly, the present invention provides for the replacement of the specificity of the reactive site of the human standard mechanism inhibitor by residues which confer specificity to subtilisin. In this way, an unmodified human standard mechanism inhibitor is used as a "backbone", and the reactive site of the human standard mechanism inhibitor is recombinantly engineered to be more specific for subtilisin, thus forming an inhibitor specific for subtilisin without undesirable immunogenicity.

The human standard mechanism inhibitor may be modified in a variety of ways. In one embodiment, all or part of the existing reactive site, which is specific for a protease other than subtilisin, is replaced with the reactive site of a known subtilisin inhibitor, such as an ovomucoid third domain, or *Streptomyces* subtilisin inhibitor (SSI), among others. Alternatively, the reactive site of a known subtilisin inhibitor is used as a prototype, and only selected residues are modified in the human standard mechanism inhibitor in order to confer proper specificity for subtilisin, as outlined below. In this embodiment, some of the original residues of the reactive site of the human standard mechanism inhibitor remain unchanged, and some are replaced by the residues present in the known subtilisin inhibitors. However, the new residues are sufficient to confer the appropriate subtilisin specificity to the human standard mechanism inhibitor.

Thus, the novel subtilisin inhibitors of the present invention comprise a modified human standard mechanism inhibitor, which has been recombinantly engineered to substitute at least one amino acid residue of the reactive site with a different amino acid to increase the specificity of the modified human standard mechanism inhibitor towards subtilisin.

By "subtilisin" herein is meant any number of serine endoproteases or protease variant of the "subtilase" family, preferably but not limited to alkaline serine endoprotease, produced by gram-positive bacteria or fungi, known in the art as subtilisin (Siezen et al., Protein Eng. 4:719-737 (1991)).. A "target subtilisin" is a subtilisin which is inhibited by a modified subtilisin inhibitor of the present invention. Examples of suitable subtilisins include, but are not limited to, the subtilisins from *B. amyloliquefaciens*, *B. licheniformis*, *B. subtilis*, and *B. lentus*, including subtilisin DY, subtilisin BPN', subtilisin Carsberg, subtilisin 168, subtilisin amylosaccharitius, mesentericopeptidase, subtilisin thermitase, and proteinase K.

Also included within the definition of subtilisin are variant or mutant subtilisins. A large number of variant subtilisins are known in the art. For example, variants have been made which alter a number of characteristics of the enzyme, including pH stability and optima, thermostability, substrate specificity, and oxidative resistance. See for example Estell, J. Biotechnology 26:25-29 (1993); and Wells and Estell, Trends in Biochem. Sci. 13(8): 291-297 (1988). In general, a great deal is known about the substrate specificity of variants of subtilisin.

The substrate specificity of any given subtilisin variant may be determined using a variety of synthetic substrates.

Specifically, subtilisin variants which may be inhibited with the novel inhibitors of the present invention include, but are not limited to, variants at the residue positions equivalent to the following positions in *B. amyloliquifaciens*: Tyr6, Ser9, Ile11, Lys12, Gn19, Tyr 21, Thr22, Ser24, Asn25, Asp32, Ser33, Asp36, Ser37, Ser39, Ala45, Gly46, Ala48, Ser49, Met50, Ser53, Glu54, Thr55, Asn56, Pro57, Phe58, His64, His67, Thr71, Asn76, Asn77, Ser87, Ser89, Lys94, Val95, Leu96, Gly97, Asp99, Ser101, Gly102, Glu103, Tyr104, Ile107, Gly110, Ile111, Ile115, Asp120, Val121, Ile122, Met124, Leu126, Gly127, Gly128, Pro129, Leu135, Gly131, Leu135, Asp140, Ala152, Ala153, Gly154, Asn155, Glu156, Gly157, Thr158, Ser159, Gly160, Ser161, Ser162, Ser163, Thr164, Val165, Gly166, Pro168, Gly169, Lys170, Tyr171, Pro172, Ile175, Val180, Ser182, Arg186, Ala187, Phe189, Ser191, Pro194, Glu195, Asp197, Met199, Ser204, Lys213, Tyr214, Gly215, Tyr217, Asn218, Ser221, Met222, His226, Ile234, Leu235, Ser236, Lys237, His238, Trp241, Ser260, Phe261, Tyr262, Lys265, Ile268, or Gln275. See U.S. Reissue 34,606; WO 89/06279; WO 92/10755; and 0 549 675 B1, expressly incorporated herein by reference.

By "reactive site" herein is meant the region of the inhibitor with which the enzyme interacts during inhibition. That is, the reactive site of the inhibitor occupies the active site of its target enzyme. As is appreciated in the art, the exact location and size of the reactive site will vary with the inhibitor and protease to be inhibited. The reactive sites for a number of human standard mechanism inhibitors and subtilisin reactive sites for known subtilisin inhibitors are shown in the figures.

The reactive site generally comprises residues P10 to P10', where the P1-P1' residues are the ones which occupy the S1-S1' position of the enzyme's substrate (or inhibitor) binding site. That is, at least one amino acid within this region is altered, although, as outlined below, in preferred embodiments, more than one amino acid is substituted. Preferably, the residue or residues to be

altered range between the P9 and P5′ residues which occupy the S9 to S5′ position of the enzyme binding site. In alternative embodiments, as outlined below, the residues from a smaller area of the reactive site are altered; for example, between P6 to P5′, or P5 to P3′, although it will be appreciated that any subsection of this area may be altered.

In a preferred embodiment, when the reactive site of the human standard mechanism inhibitor contains cysteine residues which form disulfide bridges, the cysteine residues are not altered, to maximize the inhibitor stability. It should be noted that in some cases, the subtilisin reactive site which is to replace the human standard mechanism inhibitor reactive site also contains cysteine residues at the identical positions. Thus, in the case where the entire reactive site of the human standard mechanism inhibitor is replaced by a subtilisin reactive site, the cysteine residues remain in the same location. In alternative embodiments, the cysteine residues may be removed.

By "subtilisin reactive site" herein is meant a reactive site, as defined above, in a known subtilisin inhibitor that is either specific for a target subtilisin or which inhibits a target subtilisin to a significant degree. That is, a subtilisin reactive site that inhibits only subtilisins is a subtilisin reactive site which is specific for subtilisin. Alternatively, a subtilisin reactive site in a known subtilisin inhibitor may significantly inhibit subtilisin, i.e. exhibit a low $K_i$ as is more fully discussed below, as well as inhibit other serine proteases.

Known subtilisin inhibitors with the P10 to P10' subtilisin reactive sites and P1-P1' sites shown are depicted in the Figures. Known subtilisin inhibitors with subtilisin reactive sites suitable for the use in the present invention include, but are not limited to, the third domains of avian ovomucoids and the *Streptomyces* subtilisin inhibitors (SSI). Included within the definition of SSI are the SSI-like inhibitors. Additionally, the sequence of a known subtilisin inhibitor from

*Bombyx mori* (silkworm) has been determined, with a $K_I$ of approximately $10^{-11}$ to $10^{-13}$ (see Eguchi et al., J. Biochem. 115:881-884 (1994)), the sequence of which is shown in Figure 13. While the reactive site of the silkworm subtilisin inhibitor has not yet been reported, well known techniques may be used to determine the reactive site (see Ozawa et al., J. Biol. Chem. 241:3955-3961 (1966) and Terabe et al., J. Biochem. 116:1156-1163 (1994), both of which are hereby incorporated by reference).

In addition to the ovomucoid third domains and the SSI subtilisin reactive sites, subtilisin reactive sites from other subtilisin inhibitors may be used. Generally speaking, a subtilisin with a suitable subtilisin reactive site has a $K_I$ for a subtilisin of $10^{-6}$ or less, with $10^{-6}$ to $10^{-14}$ being suitable, and $10^{-8}$ to $10^{-12}$ being preferred. Thus, for example, the reactive sites from the potato family of inhibitors, and the Bowman-Birk family of inhibitors may be used if the $K_I$s are suitable (see Bode et al., Eur. J. Biochem. 204:433-451 (1992), incorporated by reference).

In a preferred embodiment, the subtilisin reactive sites from avian ovomucoid third domains are used in the present invention. Avian ovomucoids generally comprise three domains, the last of which are well known as potent subtilisin inhibitors. The most commonly used ovomucoid for this purpose is the turkey ovomucoid third domain (abbreviated TOM or OMTKY.3), which has an approximate $K_I$ towards subtilisin of $10^{-9}$ to $10^{-10}$. The subtilisin reactive site of OMTKY.3 is a particularly preferred subtilisin reactive site in the invention.

The sequence of a large number of ovomucoid third domains are shown in Figure 5, with the P10 to P10′ and P1-P1′ sites shown. In a preferred embodiment, the reactive site residues to be used comprise the P6 to P5′ residues. Any of the subtilisin reactive sites of the ovomucoid third domain proteins depicted in Figure 5 may be used in the present invention. As outlined

above, the ovomucoid third domains have cysteine residues in the P3 and P6' positions, which are generally conserved between the human standard mechanism inhibitors and the known subtilisin inhibitors.

In a preferred embodiment, the subtilisin reactive sites from *Streptomyces* subtilisin inhibitors (SSI) and "SSI-like" inhibitors, (collectively "SSI inhibitors") are used in the present invention. These include, but are not limited to, SSI from *S. albogriseolus* (see Obata et al., J. Biochem. 105:367-371 (1989)) which has an approximate $K_i$ for subtilisin in the range of $10^{-12}$; the SSI-like (SIL) protein SIL1 of Kojima et al., Biochim. Biophys. Acta 1207:120-125 (1994); the SIL2, SIL3 and SIL4 inhibitors of Taguchi et al., Eur. J. Biochem. 220:911-919 (1994); the *Streptomyces lividans* inhibitor (SLPI, not to be confused with the SLPI human standard mechanism inhibitor) of Ueda et al., J. Biochem. 112:204-211 (1992); see also STI-1 of Strickler et al., J. Biol. Chem. 267(5):3236-3241 (1992); STI-2 inhibitor of Stricker et al., supra; plasminostreptin (PSN) from *S. antifibrinolyticus*, Kakinuma et al., J. Biol. Chem. 252(5):1529-1537 (1978); API-2c' inhibitor from *S. griseoincarnatus*, Suzuki et al., Agric. Biol. Chem 45(3):629-634 (1981); the SIL8, SIL10, SIL13, and SIL14 inhibitors of Terabe et al., J. Biochem. 116:1156-1163 (1994), and the anticoagulant inhibitor STVCI, from *Streptoverticillium cinnamoneum* (J. Biochem. 115:752-761 (1994), all of which are hereby expressly incorporated by reference. Furthermore, Taguchi et al. Applied Environ. Microbio. 59(12):4338-4341 (1993), also expressly incorporated by reference, describe a number of SIL proteins from a variety of *Streptomyces* species. As for the ovomucoid third domains, the preferred reactive site residues of the SSI inhibitors for substitution into human standard mechanism inhibitors comprise the P6 to P5' residues. In addition, most, if not all, of the SSI inhibitors have cysteines in the P3 position, which allows the cysteine at this position to be kept constant.

In addition, the reactive site of chelonianin is used as a subtilisin reactive site. The second domain of chelonianin, a two domain inhibitor of the red sea turtle, is highly homologous to MPI (see Seemuller et al., FEBS Lett. 199(1):43-48 (1986); and Kato et al., Fed. Proc. 38:832 (1979)). Chelonianin has been characterized as a subtilisin inhibitor. The P10 to P10' subtilisin reactive site of chelonianin is shown in Figure 7A, with the P1-P1' junction depicted. In a preferred embodiment, the chelonianin subtilisin reactive site replaces the reactive site of either domain of the MPI backbone.

In an additional embodiment, the reactive site of domain 5 of rat agrin is used as a subtilisin reactive site. The sequence of rat agrin domain 5 is shown in Figure 7B, with the P10 to P10' subtilisin reactive site depicted. Rat agrin domain 5 has recently been shown to be a subtilisin inhibitor, with a $K_i$ in the $10^{-5}$ to $10^{-6}$ range (data not shown).

Also included within the definition of "subtilisin reactive site" are modified, i.e. non-naturally occurring, reactive sites from known subtilisin inhibitors. That is, the subtilisin reactive sites of known subtilisin inhibitors may be recombinantly engineered as well to contain mutations within the subtilisin reactive site which are then exchanged with the reactive sites of human standard mechanism inhibitors to create the modified subtilisin inhibitors of the invention. For example, the P1 residue of turkey ovomucoid third domain has been changed from Leu18 to Glu18 to produce a potent inhibitor of Glu-specific *Streptomyces griseus* protease (Komiyama et al. J. Biol. Chem. 266(17):10727-10730 (1991)). In addition, mutants have been generated at the P1 position in SSI and found to retain specificity for subtilisin BPN' (Kojima et al., J. Biochem. 109:377-382 (1991), expressly incorporated by reference herein).

Thus, in a preferred embodiment, a mutant of the SSI reactive site may be used, as is depicted in Figure 11. This mutant reactive site is based on the SSI

reactive site but contains Ala at P6, Ala at P5, Ala at P4, Leu at P1, and Arg at P3'.

As will be appreciated by those in the art, residues outside the reactive site may also interact with the enzyme, as is discussed below. These residues may also be substituted to optimize the interaction of the inhibitor and the subtilisin.

By "standard mechanism inhibitor" herein is meant an inhibitor which functions according to the standard mechanism outlined in Laskowski et al., Ann. Rev. Biochem. 49:593-626 (1980), hereby incorporated by reference. This mechanism can be written as

$$E + I \approx L \approx C \approx X \approx L^* \approx E + I^*$$

Where E is the enzyme, I* and I are modified inhibitor (reactive site peptide bond hydrolyzed) and virgin inhibitor (reactive site peptide bond intact) respectively, L and L* are loose, noncovalent (rapidly dissociable) complexes of E with I and I*, respectively, X is the relatively long-lived intermediate in the E+I*, and C is the stable enzyme-inhibitor complex. Generally, the reactive site peptide bond is encompassed in at least one disulfide loop, which ensures that during conversion of virgin to modified inhibitor the two peptide chains cannot dissociate.

As is known in the art, standard mechanism inhibitors comprise several families of inhibitors, including the Kazal type inhibitors, the SSI type inhibitors, and the chelonianin family inhibitors. Included within the definition of human standard mechanism inhibitors, useful as backbones in the present invention, are the Kazal-type inhibitors such as PSTI, agrin and acrosin inhibitor (HUSI-II), and the chelonianin family inhibitors such as MPI. In a preferred embodiment, the definition of a standard mechanism inhibitor does not include the serpins or $\alpha_2$-macroglobulin.

By "human standard mechanism inhibitor" herein is meant one of the standard mechanism inhibitors, as defined herein, which is from a human, or homologous human sequences which may not yet have been shown to proteinase inhibitors; that is, the naturally occuring amino acid sequence derived from humans.

Included within the definition of human standard mechanism inhibitor is human pancreatic secretory trypsin inhibitor (hPSTI) and PSTI-like inhibitors and sequences such as human agrin domain 5, human agrin domain 7, human acrosin inhibitor II, human hevin and human follistatin, and human mucous proteinase inhibitor (MPI) and MPI-like inhibitors, including skin-derived antileukoproteinase (SKALP; also known as human elafin precursor or ESI) and human epididymius protease inhibitor, HE4.

It should be understood that some of the unmodified human standard mechanism inhibitors may inhibit subtilisin without further modification. Thus, for example, unmodified hPSTI exhibits a $K_I$ of roughly $10^{-7}$ towards subtilisin. Accordingly, as will be more fully described below, the modified human standard mechanism inhibitors of the invention, i.e. the modified subtilisin inhibitors, have decreased $K_I$s in relation to the unmodified human standard mechanism inhibitors.

It should be understood that some of the human standard mechanism inhibitors are comprised of a number of domains, not all of which may be relevant to the present invention. For example, as outlined below, agrin is a multidomain protein. The rat agrin sequence demonstrates nine tandem protease inhibitor domains, followed by a few unrelated domains. The sequence of human agrin protease inhibitor domain 7 is known, as depicted in Figure 3; however, other domains of agrin may also be used, if appropriate.

Similarly, as outlined below, MPI contains two domains, either or both of which may be engineered as taught herein. HE4 also appears to be two domains.

It is important to note that the exact "start" of the protease inhibitor domain of a multidomain protein, including SKALP, MPI, agrin and the ovomucoid domains, may vary, as one skilled in the art will appreciate. That is, while the sequences depicted herein show specific N- and C-termini, one or more residues may be added or deleted without significantly affecting the present invention. In some cases, additional residues from the "linker" sequences between domains, or parts of other domains, may be included at the N- or C-terminus of a modified subtilisin inhibitor of the invention, as will be determined by those in the art. In fact, it may be possible to include an entire additional domain, or more, in the novel subtilisin inhibitors of the invention, although this is generally not preferred. That is, when the human standard mechanism inhibitor is a multidomain protein, two or more of the domains may be included, with at least one of these domains being recombinantly engineered to have an increased specificity for subtilisin, as outlined herein. In some cases, it may be possible to engineer two or more domains for subtilisin. Thus, for example, MPI has two domains, depicted as "1" and "2" herein, both or either of which may be altered for increased subtilisin specificity. Similarly, HE4 appears to be two domains, either or both of which may be engineered as well.

In addition, the exact N- or C-termini of the inhibitors of the present invention may also vary somewhat, due to putative proteolytic processing. Thus, for example, Szardenings et al., Protein Engineering 8(1):45-52 (1995), depicts the N-terminus of hPSTI as DSLGREA---. The Szardenings sequence was determined on the basis of protein sequencing. The N-terminus of the hPSTI of the examples has five additional N-terminal amino acids, GNTGADSLGREA--- ; this sequence was obtained from the DNA sequence. It is possible to make inhibitors starting at any one of these positions, and all are included within the

definition of hPSTI. This difference would appear to result from post-translational proteolytic processing.

Thus, the N- and C-termini of the novel subtilisin inhibitors of the invention may vary from those actually depicted herein, but the addition or deletion of a few residues, or even an entire domain, is not determinative.

In a preferred embodiment, the human standard mechanism inhibitor is hPSTI, and has the sequence depicted in Figure 3. The depicted reactive site to be altered comprises residues P6 to P5'. In particular, the residues which may be altered include Leu13 (P6), Asn14 (P5), Gly15 (P4), Thr17 (P2), Lys18 (P1), Ile19 (P1'), Tyr20 (P2'), Asp21 (P3'), and Val23 (P5') (amino acid numbering assumes DSLGREA as the N-teminus). In a preferred embodiment, Cys16 (P3) is not altered, since it forms a disulfide bridge within hPSTI which is putatively important for protein stability. Similarly, Cys24 (P6') is also not altered for the same reason, although in some embodiments these residues may be modified.

In addition, residues P5 to P9 may also be changed, since computer modelling suggests steric clashes with subtilisin at subtilisin residues 99-103. Further, Thr30 and Pro32 may also be replaced, since computer modelling suggests they closely approach subtilisin residues 155 and 156.

In a preferred embodiment, the human standard mechanism inhibitor is acrosin inhibitor II (GenBank accession number P20155). Acrosin inhibitor II has the sequence depicted in Figure 3, which portrays the putative P1-P1' position and P10-P10' reactive site. The reactive site of acrosin inhibitor II was determined by alignment with the hPSTI sequence, and has been reported to inhibit acrosin and trypsin (see Moeritz et al., Gene 123:277-281 (1993)).

In a preferred embodiment, the residues to be altered include the P6 to P5' residues, namely Leu16 (P6), Pro17 (P5), Gly18 (P4), Pro20 (P2), Arg21 (P1), His22 (P1'), Phe23 (P2'), Asn24 (P3'), Pro25 (P4') and Val26 (P5') (numbering based on GLFSK-- as the N-terminus). In a preferred embodiment, Cys19 (P3) and Cys27 (P6') are not altered, due to putative disulfide bridge formation.

In a preferred embodiment, the human standard mechanism inhibitor is a region of human agrin homologous to the PSTI-like domain 7 of rat agrin (see Rupp et al., J. Neurosci. 12:3535-3544 (1992); GenBank accession number S44195). The human domain 7 is depicted in Figure 3, and the P1-P1' and putative reactive site (P10 to P10') identified. These sites were identified as for acrosin inhibitor II, by alignment with the hPSTI sequence. In a preferred embodiment, the residues to be altered include, Val20 (P7), Pro22 (P5), Arg23 (P4), Glu25 (P2), His26 (P1), Pro27 (P1'), Pro28 (P2'), His29 (P3'), Gly30 (P4'), Pro31 (P5'), Val32 (P6'), Gly33 (P8'), Ser34 (P9'), Asp35 (P10'), and Gly36 (P11') (note numbering is from the sequence shown in Figure 3 shown for domain 7 only). In a preferred embodiment, the cysteines at positions P3, P6, P8 and P7' are not altered, although they may be if desired.

In a preferred embodiment, the human standard mechanism inhibitor is human hevin protein, the sequence of which is depicted in Figure 3 with the putative reactive site depicted (see Girard et al., Immunity 2:113-123 (1995); Genbank accession number X82157 and X86693). In a preferred embodiment, the residues to be altered include, Val29 (P10), Gln31 (P8), Asp32 (P7), Pro33 (P6), Val34 (P5), Thr35 (P4), Pro37 (P2), Pro38 (P1), Thr35(P1'), Lys40 (P2'), Pro41 (P3'), Leu42 (P4'), Asp43 (P5'), Gln44 (P6'), Val45 (P7'), Gly47 (P9'), and Thr48 (P10') (note numbering is from the sequence shown in Figure 3 only and is not necessarily from the N-terminus of hevin). In a preferred

embodiment, the cysteines at positions P3, P9, and P8' are not altered, although they may be if desired.

In a preferred embodiment, the human standard mechanism inhibitor is human follistatin, the sequence of which is depicted in Figure3 with the putative reactive site depicted (Genbank accession numbers M19480 and J03771). In a preferred embodiment, the residues to be altered include, Val29 (P8), Ala31 (P6), Pro32 (P5), Asp33 (P4), Ser35 (P2), Asn36 (P1), Ile37 (P1'), Thr38 (P2'), Trp39 (P3'), Lys40 (P4'), Gly41 (P5'), Pro42 (P6'), Val43 (P7'), Gly45 (P9'), and Leu46 (P10') (note numbering is from the sequence shown in Figure 3 only and is not necessarily from the N-terminus of follistatin). In a preferred embodiment, the cysteines at positions P3, P7, P9, and P8' are not altered, although they may be if desired.

In a preferred embodiment, the human standard mechanism inhibitor is MPI. Mucous protease inhibitors have been identified in humans and called a variety of names, including secretory leukocyte protease inhibitor (SLPI), bronchial mucous inhibitor (BMI), cervical uteri secretion inhibitor (CUSI), bronchial leukocyte protease inhibitor (BLPI), and human seminal secretion inhibitor (HUSI-I, also called antileukoprotease). To date, it appears that all of these are derived from a single polypeptide with possible different proteolytic processing at the N-terminus (see Grutter et al., EMBO J. 7(2):345-351 (1988); but see Klasen et al., Biochem. Biophys. Res. Comm. 128(1):285-289 (1985)).

The sequence of MPI (identical to HUSI-I) is depicted in Figure 4, with the P1-P1′ sites and the P10-P10′ reactive sites shown. As outlined above, MPI appears to have two domains, arbitrarily designated in Figure 4 as the "1" domain and the "2" domain. Either or both of the domains find use in the present invention. That is, in one embodiment the reactive site of only a single domain is altered, while the other domain is left unchanged. Alternatively, one

domain is altered and the other is deleted. In yet another embodiment, both domains are altered.

In the "1" domain, the residues that will be altered include Ser15, Ala15, Gln17, Leu19, Arg20, Tyr21, Lys22, Lys23, Pro24, (reactive site inferred from domain 2) and Glu25, with Cys18 and Cys26 remaining unchanged in a preferred embodiment.

In the "2" domain, the residues to be altered include Tyr68 (P5), Gly69 (P4), Gln70 (P3), Leu72 (P1), Met73 (P1'), Leu74 (P2'), Asn75 (P3'), Pro76 (P4'), Pro77 (P5'), Gln78 (P6') and Phe79 (P7'), with Cys71 (P2) and Cys80 (P8') remaining unchanged in a preferred embodiment.

In a preferred embodiment, the human standard mechanism inhibitor is a domain of human skin-derived antileukoproteinase (SKALP). SKALP has also been called human elafin precursor (ELAFIN) and residues 68 to 117 of SKALP comprise the protease inhibitor domain that exhibits a 40% homology to MPI. The sequence of one domain of SKALP is shown in Figure 4, with the putative reactive site P10 to P10' and the P1-P1' junction displayed.

In a preferred embodiment, the residues to be altered include the P6 to P5' residues, namely Ile14 (P6), Leu15 (P5), Ile16 (P4), Arg17 (P3), Ala19 (P1), Met20 (P1'), Leu21 (P2'), Asn22 (P3'), Pro23 (P4'), Pro24 (P5'), Asn25 (P6') and Arg26 (P7' (note numbering is from sequence depicted in Figure 4 and is not from the N-terminus of the complete SKALP protein). In a preferred embodiment, Cys17 (P2) and Cys26 (P8') are not altered, due to putative disulfide bridge formation.

In an additional embodiment, the human standard mechanism inhibitor is a human epididymis protease inhibitor (HE4; see Kirchhoff et al., Biol. Reprod.

45:350-357 (1991). The sequence of HE4 is depicted in Figure 4, with the putative P10-P10' reactive site and the P1-P1' site shown. It should be understood that the precise reactive site of either domain of HE4 (nor MPI-1 and SKALP) is not currently known, although it may be determined by those in the art as outlined above for silkworm, and thus the site depicted in Figure 4 is one of several possibilities. HE4 is also putatively comprises two domains, with the "1" and "2" domains depicted in Figure 4.

As for the other human standard mechanism inhibitors, the reactive site of HE4 may be altered. In the "1" domain, the residues to be altered include the P6 to P6' residues, namely Ala15 (P6), Asp16 (P5), Gln17 (P4), Asn18 (P3), Thr20 (P1), Gln21 (P1'), Glu22 (P2'), Val24 (P4'), Ser25 (P5') and Asp26 (P6'). In a preferred embodiment, Cys19 (P2) and Cys23 (P3') are not altered, due to putative disulfide bridge formation. It should be noted that the actual amino acids at the P6-P6' positions may be different if the P1-P1' junction is shifted.

In the "2" domain, the residues to be altered include the P6 to P6' residues, namely Gln64 (P6), Leu65 (P5), Gly66 (P4), Leu67 (P3), Arg69 (P1), Asp70 (P1'), Gln71 (P2'), Gln73 (P4'), Val74 (P5') and Asp75 (P6'). In a preferred embodiment, Cys68 (P2) and Cys72 (P3') are not altered, due to putative disulfide bridge formation.

Once the unmodified human standard mechanism inhibitor to be used as the backbone for the modified subtilisin inhibitors of the present invention is selected, a reactive site which is particularly effective against subtilisin ("subtilisin reactive site") is constructed to replace the human standard mechanism inhibitor reactive site. This engineering of the human standard mechanism inhibitor to contain a subtilisin reactive site may be done in several ways. In one embodiment, part or all of the human standard mechanism inhibitor reactive site may be replaced. That is, the P10 to P10' residues may be

excised and replaced with the P10 to P10′ residues of a subtilisin reactive site from a known subtilisin inhibitor. Alternatively, in a preferred embodiment, direct replacement of less than the entire reactive site may be accomplished. For example, the P6 to P5′ residues may be exchanged. In a preferred embodiment, there is no "replacement" per se; rather, only selected residues within the reactive site of the human standard mechanism inhibitor are substituted by the appropriate residues at the equivalent position in the subtilisin reactive site, thus using the subtilisin reactive site from a known subtilisin inhibitor as a model or prototype.

In a preferred embodiment, the reactive site residues of the unmodified human standard mechanism inhibitor are replaced with the reactive site residues of a known subtilisin inhibitor, to generate a modified subtilisin inhibitor. As discussed above, this may be done using the entire reactive site (P10 to P10′), or portions of the reactive site (preferably P6 to P5′).

For example, in a preferred embodiment, the P6 to P5' residues of hPSTI are replaced with the P6 to P5' residues of an ovomucoid third domain protein, preferably OMTKY.3, or the P6 to P5' residues of an SSI protein, preferably the SSI from *S. albogriseolus*.

Preferred standard mechanism inhibitors/subtilisin reactive site pairs include hPSTI/OMTKY3, hPSTI/ovomucoid, hPSTI/SSI, MPI-1 or 2/red sea turtle inhibitor, MPI/OMTKY3, and MPI/SSI, although those in the art will appreciate that any combination of backbone and subtilisin reactive site may be used.

In an alternate embodiment, only selected residues within the human standard mechanism inhibitor reactive site are altered. This is done using a subtilisin reactive site as a model or prototype. Thus, for example, at least one amino

acid of the human standard mechanism inhibitor reactive site is substituted with the amino acid at the equivalent position in the subtilisin reactive site. For example, the lysine at the P1 position of hPSTI may be substituted with a leucine, which is the P1 residue of OMTKY.3, or a methionine, which is the P1 position of SSI. In other embodiments, more than two, three, four or five residues may be altered, with at least three being preferred and at least five being particularly preferred.

In order to select which residues should be substituted, one of two methods may be employed. If the structure of the inhibitor and/or reactive site is known, modeling may be done as is known in the art. Alternatively, each amino acid residue of the reactive site may be changed to each of the other amino acids, or a subset of the amino acids, and the resulting mutants screened, as is known in the art. Routine experimental techniques are available to easily mutate and screen the backbones. In general, the P1, P2 and P4 residues of the human standard mechanism inhibitor are altered, since subtilisin has distinct binding pockets at the S1, S2 and S4 sites corresponding to these sites. Thus, these three residues within the human standard mechanism may be altered first.

In addition, the characteristics of the target subtilisin may be useful in constructing a modified subtilisin inhibitor. That is, inhibitors can be designed to mimic the substrate specificity of the target subtilisin, a process that is known in the art. See for example, Lu et al., J. Biol. Chem. 268(20):14583-14585 (1993); Komiyama et al., J. Biol. Chem. 266(17):10727 (1991).

The modified subtilisin inhibitors of the present invention are constructed such that the kinetic parameters of the modified inhibitor are altered relative to the unmodified inhibitor. This may be quantified in several ways.

In a preferred embodiment, the dissociation constant for inhibition of a subtilisin by the modified subtilisin inhibitor is lower than the dissociation constant of the unmodified human standard mechanism inhibitor. As noted above, some unmodified human standard mechanism inhibitors exhibit some level of inhibition of subtilisin, although it is generally not significant enough for practical applications. Accordingly, what is important in the present invention is that the dissociation constant for subtilisin decreases as a result of the engineering of the reactive site to be more specific for subtilisin.

Accordingly, in a preferred embodiment, the modified human standard mechanism inhibitors of the invention have dissociation constants for at least one subtilisin that is decreased by at least a factor of 10 relative to the unmodified human standard mechanism inhibitor. In a preferred embodiment, the dissociation constant is decreased by a factor of 100, or $10^{-2}$, with $10^{-3}$, $10^{-4}$, $10^{-5}$ and $10^{-6}$ also possible. A decrease of at least $10^{-2}$ is preferred.

By the term "inhibition constant" or "dissociation constant" or grammatical equivalents herein is meant the equilibrium dissociation constant for the reversible association of inhibitor with enzyme. The dissociation constant for inhibition of a protease with the inhibitor is generally referred to by those in the art as $K_i$. The dissociation constants are defined and determined as below.

The determination of dissociation constants is known in the art. For example, for reversible inhibition reactions such as those of the present invention, the reaction scheme is as follows:

Equation 1

$$E + I = E \cdot I$$

The enzyme and the inhibitor combine to give an enzyme-inhibitor complex, E·I. This step is assumed to be rapid and reversible, with no chemical changes taking place; the enzyme and the inhibitor are held together by non-covalent forces. In this reaction, $k_1$ is the second order rate constant for the formation of the E·I reversible complex, i.e. the "on" rate, referred to herein as "$k_{on}$". $k_2$ is the first order rate constant for the disassociation of the reversible E·I complex, i.e. the "off" rate, referred to herein as "$k_{off}$". In this reaction, $K_I = k_2/k_1$.

The measurement of the equilibrium constant $K_I$ proceeds according to techniques well known in the art, as described in the examples. For example, assays generally use synthetic chromogenic or fluorogenic substrates. Particularly preferred are methods for the determination of $K_I$s for tight-binding inhibitors. The $K_I$ may be determined in several ways, as outlined below.

In one embodiment, the $K_I$ is determined as a ratio of rate constants, as outlined in Tonomura et al., in Protein Protease Inhibitor - The Case of Streptomyces Subtilisin Inhibitor, Hiromi et al., ed., Elseview Science Publishers B.V., 1985, chapter 9, particularly pages 297-299. Briefly, the $K_I$ is determined as the ratio of the $k_{on}$ and $k_{off}$ rates, which may be determined either directly or graphically. This method is outlined in Example 2.

The $K_I$ for competitive binding inhibitors may also be determined from the following calculation:

Equation 2

$$[I]_o/\{1-(v_i/v_o)\} = \{K_I/(v_i/v_o)\} + [E]_o$$

wherein

$[I]_o$ is total inhibitor concentration, $[E]_o$ is the total enzyme concentration, $v_o$ is the rate of substrate hydrolysis in the absence of inhibitor, and $v_i$ is the rate in

the presence of competitive inhibitor. A plot of $[I]_o/\{1-(v_i/v_o)\}$ vs. $v_o/v_i$ gives a slope equal to $K_I(1 + [S]/K_m)$, which is the apparent $K_I$, with an x-intercept of $[E]_o$.

It is to be understood that dissociation constants are a particularly useful way of quantifying the efficiency of an enzyme with a particular substrate or inhibitor, and are frequently used in the art as such. If an inhibitor exhibits a very low $K_I$, it is an efficient inhibitor. Accordingly, in a preferred embodiment, the $K_I$ ranges from $10^{-2}$ to $10^{-12}$, with $10^{-4}$ to $10^{-10}$ being preferred, and $10^{-7}$ to $10^{-10}$ being particularly preferred. However, the $K_I$ should not be so low as to prevent the release of the inhibitor from the subtilisin enzyme upon dilution, as is outlined below. Thus, generally, $K_I$s of less than $10^{-10}$ are not preferred, although $K_I$s in this range are acceptable if the $k_{off}$ is fast enough.

In addition, in a preferred embodiment, the modified subtilisin inhibitors of the invention have dissociation constants which are low enough to allow adequate and efficient inhibition of the subtilisin in a concentrated form or in a detergent formulation, yet have $k_{off}$ rates which allow fast release of the inhibitor upon dilution of the detergent or formulation.

Accordingly, in a preferred embodiment, the modified subtilisin inhibitors preferably have $k_{off}$ rates which reflect fast half times of dissociation (half time of dissociation = $0.693/k_{off}$). Preferred half times of association range from about 15 seconds to about 300 seconds, or $k_{off}$s ranging from about 0.0462 to about 0.00231. However, as will be appreciated by those in the art, there is a balance between fast half times of dissociation and the $K_I$. The $k_1$ and $k_2$, ie. $k_{on}$ and $k_{off}$ constants are determined using well known techniques, as is generally described above and in the Examples.

In one embodiment, when the inhibitor is to be used as part of an enzyme and detergent composition, the ratio of the dissociation constant to the protease concentration is in the range from 0.006 to 6.

The $K_I$ may be also be determined by measuring the $I_{50}$ (or $IC_{50}$). Generally, $IC_{50}$ is the concentration of inhibitor at which the enzyme reaction velocity is 50% of the uninhibited reaction. When the enzyme is preincubated with a tight-binding inhibitor, the $K_I$ may be determined via Equation 3:

$$IC_{50} = 1/2E_t + K_I$$

where $E_t$ is total enzyme concentration in molar. $IC_{50}$s are determined according to Cha et al., Biochem. Pharmacology 24:2187-2197 (1975); see also Cha et al. Biochem. Pharmacology 24:2177-2185 (1975), both of which are incorporated by reference.

As for $K_I$, $IC_{50}$ measurements are a useful way of quantifying the efficiency of an inhibitor. The lower the $IC_{50}$, the more efficient the inhibitor.

The assays for the determination of kinetic parameters are done as is known in the art, as is shown in the Examples. For example, the reactions are usually conducted at the pH optimum of the subtilisin enzyme, in the presence of small concentrations of Triton X-100 or other detergents, as subtilisin tends to stick to disposable cuvettes and microcentrifuge.

In a preferred embodiment, the human standard mechanism inhibitor is hPSTI, and the subtilisin reactive site is from an ovomucoid third domain. A particularly preferred subtilisin reactive site is the turkey ovomucoid third domain subtilisin reactive site. The sequence of the modified subtilisin inhibitor of this embodiment is shown in Figure 10.

In a preferred embodiment, the human standard mechanism inhibitor is hPSTI, and the subtilisin reactive site is from a SSI inhibitor. A particularly preferred subtilisin reactive site is the SSI subtilisin reactive site from *S. albogriseolus*. The sequence of the modified subtilisin inhibitor of this embodiment is shown in Figure 9.

In a preferred embodiment, the human standard mechanism inhibitor is MPI, and the subtilisin reactive site is from either domain of chelonianin from red sea turtle. Similarly, MPI with the reactive site from OMTKY.3 or a SSI is also preferred.

Once the subtilisin reactive site and the human standard mechanism inhibitor are chosen, the nucleic acids encoding the modified subtilisin inhibitors of the present invention are made using techniques well known in the art. As used herein, "nucleic acid" may refer to either DNA or RNA, or molecules which contain both deoxy- and ribonucleotides. The nucleic acids include genomic DNA, cDNA, mRNA and oligonucleotides including sense and anti-sense nucleic acids. Generally, the nucleic acid encoding the human standard mechanism inhibitor is used as the starting material.

However, since the inhibitors of the invention are not large, it is also possible to construct the nucleic acids encoding the inhibitors *de novo*, using synthetic techniques well known in the art.

These techniques are generally outlined in the examples, and may be used when the entire reactive site, a portion of the reactive site or only selected residues within the reactive site of the human standard mechanism inhibitor are to be replaced.

As it is understood by those in the art, once the amino acid sequence of a modified subtilisin inhibitor is selected, it is possible to generate a large number of nucleic acids which code for the modified subtilisin inhibitor amino acid sequence, due to the degeneracy of the genetic code. Thus, for example, synthetic oligonucleotides may be made which use different codons than the naturally occurring or native codons, and the synthetic oligonucleotides incorporated into the modified subtilisin inhibitor nucleic acid using techniques well known in the art. In a preferred embodiment, the nucleic acid mimics as closely as possible the human naturally occuring sequence of the human standard mechanism inhibitor.

In addition, the nucleic acids encoding the modified subtilisin inhibitors of the present invention may also include additional coding regions. For example, the nucleic acids may include nucleotides which encode sequences useful in purification or identification, such as epitope or affinity tags such as M2, as outlined in the Examples. The M2 tags of the examples were all attached at the C-terminus, but N-terminal tags may be used as well (data not shown). Similarly, as outlined above, additional N- or C-terminal amino acid residues may be encoded, or additional domains.

Once the nucleic acids encoding the modified subtilisin inhibitors of the invention are made, they are cloned and expressed in a variety of host cells for additional manipulations and expression.

The expression vector and host cell will depend on the end use of the modified subtilisin inhibitor. In a preferred embodiment, the modified subtilisin inhibitors are produced in host cells, purified or concentrated as needed, and then added to subtilisin-containing compositions such as detergent formulations, as outlined below. In this embodiment, the modified subtilisin inhibitor may be produced in a variety of host cell types, as outlined below. Alternatively, the modified

subtilisin inhibitor may be co-expressed with the target subtilisin in the host cell, to eliminate the need for two fermentations. An additional advantage is that the yield of target subtilisin may also increase due to lowered autolysis.

Using the nucleic acids of the present invention which encode the modified subtilisin inhibitors of the invention, a variety of expression vectors are made. The expression vectors may be either self-replicating extrachromosomal vectors or vectors which integrate into a host genome. Particularly preferred are integration vectors for integration into subtilisin production strains, as outlined below.

Generally, these expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleic acid encoding the modified subtilisin inhibitor. "Operably linked" in this context means that the transcriptional and translational regulatory DNA is positioned relative to the coding sequence of the modified subtilisin inhibitor in such a manner that transcription is initiated. Generally, this will mean that the promoter and transcriptional initiation or start sequences are positioned 5' to the modified subtilisin inhibitor coding region. The transcriptional and translational regulatory nucleic acid will generally be appropriate to the host cell used to express the modified subtilisin inhibitor; for example, transcriptional and translational regulatory nucleic acid sequences from *Bacillus* are preferably used to express the modified subtilisin inhibitor in *Bacillus*. Numerous types of appropriate expression vectors, and suitable regulatory sequences are known in the art for a variety of host cells.

In general, the transcriptional and translational regulatory sequences may include, but are not limited to, promoter sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and stop sequences,

and enhancer or activator sequences. In a preferred embodiment, the regulatory sequences include a promoter and transcriptional start and stop sequences.

Promoter sequences encode either constitutive or inducible promoters. The promoters may be either naturally occurring promoters or hybrid promoters. Hybrid promoters, which combine elements of more than one promoter, are also known in the art, and are useful in the present invention.

In addition, the expression vector may comprise additional elements. For example, the expression vector may have two replication systems, thus allowing it to be maintained in two organisms, for example in mammalian or insect cells for expression and in a procaryotic host for cloning and amplification. Furthermore, for integrating expression vectors, the expression vector contains at least one sequence homologous to the host cell genome, and preferably two homologous sequences which flank the expression construct. The integrating vector may be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs for integrating and replicating vectors are well known in the art.

In addition, in a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

The modified subtilisin inhibitors of the present invention are produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a modified subtilisin inhibitor, under the appropriate conditions to induce or cause expression of the modified subtilisin inhibitor. The conditions appropriate for modified subtilisin inhibitor expression will vary with the choice of the expression vector and the host cell, and will be easily ascertained by one skilled in the art through routine experimentation. For example, the use of

constitutive promoters in the expression vector will require optimizing the growth and proliferation of the host cell, while the use of an inducible promoter requires the appropriate growth conditions for induction. In addition, in some embodiments, the timing of the harvest is important. For example, the baculoviral systems used in insect cell expression are lytic viruses, and thus harvest time selection can be crucial for product yield.

Appropriate host cells include yeast, bacteria, archebacteria, fungi, and insect and animal cells, including mammalian cells. Of particular interest are *Drosophila melanogaster* cells, *Saccharomyces cerevisiae* and other yeasts, *E. coli*, *Bacillus* species such as *B. subtilis* and *B. licheniformis*, SF9 cells, C129 cells, 293 cells, *Neurospora*, BHK, CHO, COS, and HeLa cells, *Streptomyces* species such as *S. lividans*, fungal strains such as *Aspergillus*, *Tricoderma* and others, immortalized mammalian myeloid, and lymphoid cell lines.

When the modified subtilisin inhibitor is produced independently from the target subtilisin, particularly preferred expression systems include *Bacillus* species such as *B. subtilis*.

When the modified subtilisin inhibitor is to be produced simultaneously with the target subtilisin, particularly preferred expression systems include *Bacillus* species such as *B. subtilis*.

In a preferred embodiment, modified subtilisin inhibitors are expressed in bacterial systems. Bacterial expression systems are well known in the art.

A suitable bacterial promoter is any nucleic acid sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of the coding sequence of modified subtilisin inhibitor into mRNA. A bacterial promoter has a transcription initiation region which is usually placed proximal

to the 5′ end of the coding sequence. This transcription initiation region typically includes an RNA polymerase binding site and a transcription initiation site. Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose and maltose, and sequences derived from biosynthetic enzymes such as tryptophan. Promoters from bacteriophage may also be used and are known in the art. In addition, synthetic promoters and hybrid promoters are also useful; for example, the *tac* promoter is a hybrid of the *trp* and *lac* promoter sequences. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription.

In addition to a functioning promoter sequence, an efficient ribosome binding site is desirable. In *E. coli*, the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon and a sequence 3-9 nucleotides in length located 3 - 11 nucleotides upstream of the initiation codon.

The expression vector may also include a signal peptide sequence that provides for secretion of the modified subtilisin inhibitor in bacteria. The signal sequence typically encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell, as is well known in the art. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria).

The bacterial expression vector may also include a selectable marker gene to allow for the selection of bacterial strains that have been transformed. Suitable selection genes include genes which render the bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin and

tetracycline. Selectable markers also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways.

These components are assembled into expression vectors. Expression vectors for bacteria are well known in the art, and include vectors for *Bacillus subtilis*, *E. coli*, *Streptococcus cremoris*, and *Streptococcus lividans*, among others.

The bacterial expression vectors are transformed into bacterial host cells using techniques well known in the art, such as calcium chloride treatment, electroporation, and others.

In one embodiment, modified subtilisin inhibitors are produced in insect cells. Expression vectors for the transformation of insect cells, and in particular, baculovirus-based expression vectors, are well known in the art.

Mammalian expression systems are also known in the art and are used in one embodiment. A mammalian promoter is any DNA sequence capable of binding mammalian RNA polymerase and initiating the downstream (3') transcription of a coding sequence for modified subtilisin inhibitor into mRNA. A promoter will have a transcription initiating region, which is usually place proximal to the 5' end of the coding sequence, and a TATA box, located 25-30 base pairs upstream of the transcription initiation site. The TATA box is thought to direct RNA polymerase II to begin RNA synthesis at the correct site. A mammalian promoter will also contain an upstream promoter element, typically located within 100 to 200 base pairs upstream of the TATA box. An upstream promoter element determines the rate at which transcription is initiated and can act in either orientation. Of particular use as mammalian promoters are the promoters from mammalian viral genes, since the viral genes are often highly expressed and have a broad host range. Examples include the SV40 early

promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter, and herpes simplex virus promoter.

Typically, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3′ to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. The 3′ terminus of the mature mRNA is formed by site-specific post-translational cleavage and polyadenylation. Examples of transcription terminator and polyadenylation signals include those derived from SV40.

The methods of introducing exogenous nucleic acid into mammalian hosts, as well as other hosts, is well known in the art, and will vary with the host cell used. Techniques include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

In a preferred embodiment, stably transfected mammalian cell lines which express modified subtilisin inhibitors are made.

In a preferred embodiment, modified subtilisin inhibitor is produced in yeast cells. Yeast expression systems are well known in the art, and include expression vectors for Saccharomyces cerevisiae, Candida albicans and C. maltosa, Hansenula polymorpha, Kluyveromyces fragilis and K. lactis, Pichia guillerimondii and P. pastoris, Schizosaccharomyces pombe, and Yarrowia lipolytica. Preferred promoter sequences for expression in yeast include the inducible GAL1,10 promoter, the promoters from alcohol dehydrogenase, enolase, glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-dehydrogenase, hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, pyruvate kinase, and the acid phosphatase gene.

Yeast selectable markers include ADE2, HIS4, LEU2, TRP1, and ALG7, which confers resistance to tunicamycin; the G418 resistance gene, which confers resistance to G418; and the CUP1 gene, which allows yeast to grow in the presence of copper ions.

A recombinant modified subtilisin inhibitor may be expressed intracellularly or secreted.

The modified subtilisin inhibitor may also be made as a fusion protein, using techniques well known in the art. The modified subtilisin inhibitor may be made as a fusion protein to increase expression. For example, the modified subtilisin inhibitor is made as a fusion protein with the subtilisin to be inhibited, preferably using a linker cleavable by the subtilisin or another enzyme present in the mixture. Once the fusion protein is secreted, the active subtilisin will cleave the linker, releasing the inhibitor, which then inhibits the subtilisin from further proteolytic activity.

In this embodiment, any number of suitable linkers may be used. Peptide linkers are preferred, with linkers of about 3 to about 10 amino acids being preferred, although longer linkers may be used. Particularly preferred are linkers of homoalanine, such as peptides of five alanine residues.

The modified subtilisin inhibitors of the invention may also be translationally coupled to another protein, as is known in the art.

As discussed above, amino acid residues of the human standard mechanism inhibitor other than those of the reactive site may also be altered. In this way, substitutional, insertional or deletional variants outside of the reactive site of human standard mechanism inhibitors are made. These modifications are ordinarily prepared in the same way as the reactive site substitutions. Thus, site

specific mutagenesis of nucleotides in the DNA encoding the modified subtilisin inhibitor is done, using cassette mutagenesis or other techniques well known in the art, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture as outlined above. Alternatively, the proteins or the nucleic acids may be prepared by in vitro amino acid or nucleic acid synthesis using established techniques. Alterations in positions outside of the reactive site typically do not alter the qualitative biological activity of the unmodified human standard mechanism inhibitor, and preferably do not increase the allergenicity of the inhibitor. In a preferred embodiment, the human standard mechanism inhibitor is altered such that any subtilisin proteolytic sites are removed. That is, the target subtilisin may cleave the human standard mechanism inhibitor outside of the reactive site, resulting in a loss of activity of the inhibitor. These secondary proteolytic sites may be altered to decrease or eliminate this proteolytic cleavage.

In a preferred embodiment, the modified subtilisin inhibitor is purified or isolated after expression. Modified subtilisin inhibitors may be isolated or purified in a variety of ways known to those skilled in the art depending on what other components are present in the sample. standard purification methods include electrophoretic, molecular, immunological and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, and chromatofocusing. For example, the modified subtilisin inhibitor may be purified using a standard antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful, particularly when the modified subtilisin inhibitor is produced with the target subtilisin. For general guidance in suitable purification techniques, see Scopes, R., Protein Purification, Springer-Verlag, NY (1982). The degree of purification necessary will vary depending on the use of the modified subtilisin inhibitor. In some instances no purification will be necessary.

In a preferred embodiment, the modified subtilisin inhibitors of the invention are co-expressed with the target subtilisin, and are subjected to whatever downstream processing the subtilisin normally receives. In this embodiment, it is expected that the inhibitor will form an enzyme/inhibitor complex with the subtilisin, and thus the two will be relatively inseparable while in a concentrated form. Thus for example, fermentation broth may be concentrated and subjected to a diafiltration and ultrafiltration step, a crystallization step, granularization or standard chromatographic procedures.

Once expressed and purified if necessary, the modified subtilisin inhibitors are useful in a number of applications.

In one embodiment, the modified subtilisin inhibitors are used as protease inhibitors in a variety of uses. For example, the inhibitors are coupled to solid supports for use as chromatography columns for the purification of subtilisin. The inhibitors may be used to titrate the active sites of subtilisin for use in quantifying the amount of subtilisin present, for example as a diagnostic tool.

In a preferred embodiment, the modified subtilisin inhibitors are incorporated into detergent formulations containing subtilisin enzymes. Detergent formulations are well known in the art. The inhibitors are combined with detergents, builders, bleach and/or fluorescent whitening agents as is known in the art for use with proteolytic enzymes. The inhibitors find particular use in formulations containing bleaching or oxidizing agents such as perborate, percarbonate or hydrogen peroxide, as the inhibitor can help stabilize the enzyme. Suitable detergents include linear alkyl benzene sulfonates, alkyl ethoxylated sulfate, sulfated linear alcohol or ethoxylated linear alcohol.

The inhibitors may be incorporated into liquid, powder or granular form. See for example U.S. Re. 34,606, EP 0 549 675 B1, U.S. Pat. Nos. 5,039,446,

5,178,789, all of which are expressly incorporated herein, and references cited therein.

When used in a detergent formulation, the modified subtilisin inhibitor is added in an amount at least about 70 percent on a molar basis of the enzyme to be stabilized, with at least about 80% to about 100% being preferred. The modified subtilisin inhibitor is added to the concentrated detergent or cleaning composition in an amount proportional to the inhibition constant ($K_I$). Thus, in a preferred embodiment, at least about 55% of the target subtilisin of the detergent is essentially bound to the inhibitor in the concentrated form, with at least about 60% being preferred and at least about 70% being particularly preferred. However, when the detergent is diluted with water or other appropriate liquid, less than 45% of the enzyme is inhibited with the modified subtilisin inhibitors of the invention. Generally this dilution is from 2 to about 10,000 or 100,000 times, with a 670 fold dilution being standard in laundry applications. In a preferred embodiment, less than about 30 % of the enzyme is inhibited upon dilution, with less than about 20% being particularly preferred. The percentage of the enzyme-inhibitor complex may be determined using Equation 1.

In a preferred embodiment, human subjects exhibit little or no immune response upon exposure to the modified subtilisin inhibitors of the present invention. That is, the modified subtilisin inhibitors have significantly reduced allergenicity as compared to known subtilisin inhibitors.

The following examples serve to more fully describe the manner of using the above-described invention, as well as to set forth the best modes contemplated for carrying out various aspects of the invention. It is understood that these examples in no way serve to limit the true scope of this invention, but rather are

presented for illustrative purposes. All references cited herein are incorporated by reference.

## EXAMPLES

### Example 1

### Cloning and Expression of Modified Subtilisin Inhibitors

Inhibitors may be produced in the appropriate organism using any well established method (for example, see Methods Enzymol. 185, 1990, the entire volume). In general, the expression system will require an appropriate transcriptional promoter, preferably inducible, upstream of the inhibitor gene with standard transcriptional termination sequences placed downstream of the gene. A ribosome binding site (Stormo, G.D., Schneider, T.K., and Gold, L.M., Nucleic Acids Res. 10, 2971-2996, 1982) or translational initiation sequences (Kozak, M., Cell 44, 283-292, 1986) will be needed for expression in prokaryotes and eukaryotes, respectively. The inhibitor gene may contain extra coding sequences at either end that simplify the gene construction or aid in the purification of the final protein product. For extracellular production, a signal sequence (von Heijne, G., Genetic Engineering 14, 1-11, 1992) may be fused to the 5' end of the inhibitor gene. The inhibitor gene may also be fused at either end to a much larger protein gene (*e.g.* to produce an inhibitor protein fused to the C-terminal end of subtilisin). The entire expression system would be introduced into an integrational or independently replicating vector suitable to the chosen organism. Readily available transformation or transfection methods could be used to introduce the expression vector into a given host. Well characterized selection methods could be used to isolate and maintain cultures that carry the desired vector. Standard protocols would be used for the growth of the organism and induction of the expression of the inhibitor gene. The inhibitor gene may be coexpressed with another protein gene (*e.g.* subtilisin) if desired.

For use in industrial products, the preferred organism for the expression of the inhibitors would be a *Bacillus* species, typically *Bacillus subtilis*. The use of various host strains, vectors, promoters, terminators, translational signals, signal sequences, and growth conditions for the production of heterologous proteins in *Bacillus subtilis* has been well documented in the literature (see Wang, L.-F., and Doi, R.H., Biotechnology 22, 63-104, 1992 and references within). Furthermore, the molecular genetics and physiology of *Bacillus subtilis* has been studied in detail (see relevant chapters in *Bacillus subtilis* and other gram-positive bacteria: biochemistry, physiology, and molecular genetics, L. Sonenshein, J.A. Hoch, and R. Losick, eds., American Society for Microbiology, D.C., 1993).

Specific examples

Inhibitor gene construction

The inhibitor genes were synthesized by using PCR (polymerase chain reaction) to amplify overlapping synthetic oligonucleotides (synthesized using an ABI 394 DNA/RNA Synthesizer). Appropriate restriction sites and M2 tag purification sequences were added to the inhibitor gene sequences to simplify cloning and protein purification, respectively (see Figures). The DNA sequence was verified using a 373A DNA Sequencer and a Cycle Sequencing kit (ABI).

E. coli expression

The *phoA* promoter and signal sequence was amplified from *E. coli* MM294 cells by using PCR. The PCR fragment was digested with *Bgl*II and *Bss*HII and purified from a 5% polyacrylamide gel. An inhibitor gene, as a *Bss*HII-*Pst*I fragment, was ligated with the promoter fragment into the vector pKK223-3 (Pharmacia) that had been previously cleaved with *Bam*HI and *Pst*I (Figure 12). Different reactive sites were introduced by utilizing the *Avr*II and *Kpn*I restriction sites (Figures). The inhibitor expression vectors (*e.g.*

pKKphoA.PSTI.SSI.M2, Figure 2) were used to transform *E. coli* NL106 cells. Cultures were grown overnight in low phosphate media (Yuan, L., Craig, S.P., McKerrow, J.H., and Wang, C.C, *al*, J. Biol. Chem. 265, 13528-13532, 1990) for the induction of the *phoA* promoter. Significant levels of the inhibitor were produced in the extracellular medium.

*B. subtilis* expression

The replicating plasmid, pBN3 (Figure 1), is an *E. coli-B. subtilis* shuttle vector constructed from pBR322 (Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., and Boyer, H.W., Gene 2, 95-113, 1977) and pUB110 (Gryczan, T.J., Contente, S., and Dubnau, D., J. Bacteriol. 134, 318-329, 1978) with a polycloning site added using a synthetic oligonucleotide. The original *Eco*RI site in pBR322 was eliminated by filling in the site with Klenow enzyme and ligating the resultant blunt ends. The *Pst*I site in the ampicillin resistance gene of pBR322 was removed by swapping the *Fsp*I-*Aat*II fragment from the original plasmid with one from pUCBM21 (Boeringer Mannheim). The pBNppt expression vector was constructed by inserting four PCR amplified fragments into the *Eco*RI and *Hind*III sites of pBN3 (Figure 1). The *B. subtilis aprE* promoter was amplified from *B. subtilis* BG2036 cells as an *Eco*RI-*Bam*HI fragment. Sequences from the *B. amyloliquefaciens* subtilisin gene were amplified from pS4.5 (Well, J.A., Ferrari, E., Henner, D.J., Estell, D.A., and Chen, E.Y., Nucleic Acids Res. 11, 7911-7925, 1983). The signal sequence region was isolated as a *Bcl*I-*Bss*HII fragment, the pro sequence region as a *Bss*HII-*Sph*I fragment, and the terminator region as a *Sph*I-*Hind*III fragment (Figure 1). Inhibitor genes were then inserted into the pBNppt expression vector utilizing the *Bss*HII and *Bam*HI sites. The starvation method (Anagnostopoulos, C., and Spizizen, J., J. Bacteriol. 81, 741-746, 1961) was used to transformation *B. subtilis* WB600 with the inhibitor expression vectors and resistance to neomycin was used for selection. The inhibitors were produced in flasks by shaking (225 rpm) overnight at 37°C in very rich medium

(Halling, S.M., Sanchez-Anzaldo, F.J., Fukuda, R., Doi, R.H., and Meares, C.F., Biochemistry 16, 2880-2884, 1977).

Baculovirus expression

The PCR amplified genes for PSTI.M2 and for domain 5 of rat agrin were cloned into the baculovirus transfer vector pAcGP67 A (PharMingen) as *Bam*HI-*Eco*RI fragments (Figures 8 and 14). *Spodoptera frugiperda* (Sf9) were transfected and inhibitor proteins produced as specified by the manufacturer (PharMingen).

Example 2

Kinetic Characterization

Unless otherwise noted, all materials used were of reagent grade or better.

**Isolation of PSTI.SSI (PSTI scaffold with *Streptomyces* subtilisin inhibitor reactive site).**

One liter of fermentation broth of pKKphoA.PSTI.SSI..m2 was centrifuged at 10,000 rpm in a Sorvall RC-5A centrifuge for ten minutes to pellet the cells. The supernatant was removed and filtered using a 0.45$\mu$m filter. Immediately after filtering PMSF (phenylmethylsulfonylfluoride, Sigma Chemical Co., St. Louis, MO) and EDTA (ethylenedinitrolotetraacetic acid, Sigma Chemical Co., St. Louis, MO) were each added to a final concentration of 1mM. The resultant supernatant was then adjusted to pH 7.0 with a 10N NaOH solution. Since PSTI.SSI possesses a M2-flag peptide on the C-terminus, 10 ml of anti-FLAG M2 affinity resin (Eastman Kodak Co., New Haven, CT) were added to the supernatant. This suspension was gently stirred for 1h at room temperature after which the resin with PSTI.SSI.M2 attached was filtered using a 0.8$\mu$m filter (Baxter, Hayward, CA). The resin was resuspended in 5mM HEPES (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid, J.T. Baker, Phillipsburg,

NJ), 150mM NaCl, pH 7.0 and poured into a XK-16 FPLC column (Pharmacia, Upsala, Sweden). The protein was eluted from the column in the following manner: The column was first washed with 5mM HEPES, 150mM NaCl, pH 7.0 buffer until the optical density at 280 nm of the effluent reached baseline level. Next the column was washed with 5mM HEPES, 1M NaCl, pH 7.0 buffer to remove any non-specifically bound proteins. Finally the PSTI.SSI.m2 was eluted in 4 column volumes of 100mM sodium acetate, 1M NaCl, pH 4.5 buffer. Fractions containing inhibitiory activity against subtilisin A (CalBiochem, La Jolla, CA), were pooled, adjusted to pH 7.0 with 1M TRIS buffer, pH 8.0 (tris(hydroxymethyl)aminomethane. No further purification or concentration of the protein was necessary for analyses.

**Determination of Concentration of PSTI.OMTKY3.m2 and PSTI.SSI.m2 Inhibitors**

The concentrations of the PSTI.OMTKY3 and PSTI.SSI inhibitors were determined using the method of Ackermann and Potter [W.W. Ackermann and V.R. Potter, *Proc. Soc. Exp. Biol. Med.*, 72, 1 (1949)]. Briefly, 8 different concentrations of inhibitor were incubated for one hour at 5 different subtilisn A concentrations in 990$\mu$l volumes of assay buffer (100mM TRIS, 10mM $CaCl_2$, 0.05% Triton X-100, pH 8.6). After one hour the 1ml volumes were transferred to cuvettes and the residual subtilisin A activities were measured by addition of 10$\mu$l succinyl-alanyl-alanyl-prolyl-phenylalanine-p-nitroanalide [suc-AAPF-pNA; 100mg/ml in DMSO] (Bachem Bioscience, Inc, King of Prussia, PA). The rate of hydrolysis was monitored at 410 nm using a Hewlett-Packard Model 8452A Diode Array Spectrophotometer. A subtilisin A control containing no inhibitor was also carried out at each subtilisin A concentration. The resultant observed rates of hydrolysis at each inhibitor concentration were plotted against the 5 subtilisin A concentrations. The asymptotes to these curves allows for the ready acquisition of the inhibitor concentration as the y-intercepts equal $-k_s[I]$ and the slopes of the asymptotes

equal k, as described by Ackermann and Potter (W.W. Ackermann and V.R. Potter, *Proc. Soc. Exp. Biol. Med.*, **72**, 1 (1949)).

**Determination of Concentration of PSTI.m2 Inhibitor**

The concentration of PSTI.m2 was determined by a modified method of Green and Work (N.M. Green and E. Work. *Biochem. J.* **54**, 347, (1953)). A known concentration of trypsin was incubated in assay buffer with various amounts of purified PSTI.m2 for 1h at 25°C (50mM TRIS, 1mM $CaCl_2$, pH 8.0). The percent residual activity of trypsin was then measured as described in Example 2, using 10$\mu$l BAPNA (Na-benzoyl-DL-arginine p-nitroanilide; 43.5mg/ml) as substrate. The resultant residual activites were then plotted against the volumes of PSTI.m2 inhibitor used in the incubations. At a trypsin concentration which is ≈10x greater than the $K_I$, there should be a sharp break point in the titration curve. Because this inhibitor forms a 1:1 complex with trypsin, extrapolation of the asymptote from 100% activity down to this break point, such that the asymptote intersects the x-axis was found to be the concentration of inhibitor in units of millimolar.

**Determination of $K_I$ for PSTI.SSI**

By definition, $K_I = k_{off}/k_{on}$, and as such, independent measurement of both $k_{off}$ and $k_{on}$ allowed for ready calculation of $K_I$.

**Determination of $k_{off}$:**

Once having established the concentration of the inhibitor, equimolar amounts of subtilisin A and inhibitor were incubated at 25°C for 1h (1:1 complex concentration = $2.5 \times 10^{-9}$ M). A 2.5$\mu$l aliquot of this 1:1 enzyme:inhibitor complex was removed and added to 997.2$\mu$l assay buffer (100mM TRIS, 10mM $CaCl_2$, 0.05% Triton X-100, pH 8.6) containing 10$\mu$l suc-AAPF-pNA as substrate (100mg/ml in DMSO as stock solution). At the time of addition, the spectrophotometer was started and the rate of hydrolysis of substrate was monitored at 410 nm. The rate of hydrolysis was monitored until it reached a

constant value corresponding to complete dissociation of the complex (maximal hydrolysis rate of the subtilisin A control reaction). The straight line resulting from the fit to maximal rate was extrapolated back to the time axis to obtain the reciprocal ($1/k_{off}$) of the dissociation constant, $k_{off}$.

**Determination of $k_{on}$:**

Using the method described by Tonomura, et al., Interaction of SSI and proteases in solution, Chapter 9 in Protein Protease Inhibitor- The Case of Streptomyces Subtilisin Inhibitor (SSI), Ed. Hiromi et al., Elsevier, 1985, the association of subtilisin A with PSTI.SSI.m2 was followed by measuring the decrease in enzyme activity of subtilisin A and inhibitor in the presence of substrate. The association was started by mixing PSTI.SSI.m2 and subtilisin A, and after a certain time interval, an excess amount of substrate was added to the 990$\mu$l enzyme:inhibitor solution (final concentrations:subtilisin A, 0.268 nM; PSTI.SSI.m2,4.43 nM; suc-AAPF-pNA, 10$\mu$l of 100mg/ml stock in DMSO). Upon addition of the substrate, the residual enzyme activity, which is proportional to the free enzyme concentration, was monitored at 410 nm. The residual activities were plotted against the time intervals allowed for the association. This curve follows the second-order reaction, and the association rate constant, $k_{on}$, was determined according to the equation shown below:

$$\{2.303/[Io]-[Eo]\} \log\frac{[Eo]([Io]-[EI])}{[Io]([Eo]-[EI])} = k_{on}t$$

The linear portion of this plot yielded a straight line, with a slope = $\{[I_o]-[E_o]/2.303\}k_{on}$ and a y-intercept = $-\log[E_o]/[I_o]$. Therefore, $k_{on}$ was readily calculated from the slope. Results of $k_{on}$, $k_{off}$, $K_i$ determinations are shown in Table I.

**Table I: Summary of Rate C nstants and Dissociation Constants**

| Inhibitor | $k_{on}$ ($M^{-1}sec^{-1}$) | $k_{off}$ ($M^{-1}sec^{-1}$) | KI M (off/on) |
|---|---|---|---|
| PSTI.m2 | $3.542 \times 10^{5}$ | 0.054 | $1.524 \times 10^{-7}$ |
| PSTI.OMTKY.m2 | $1.1012 \times 10^{7}$ | 0.0044 | $4.348 \times 10^{-10}$ |
| PSTI.SSI.m2 | $1.317 \times 10^{8}$ | 0.001 | $7.593 \times 10^{-12}$ |

**Determination of $K_I$ for PSTI.SSI using $IC_{50}$ Values**

As discussed by Cha (S. Cha *Biochem. Pharmacology* **24,** 2177, (1975)), under experimental conditions outlined above (pre-incubation of the inhibitor with the enzyme), the $IC_{50}$ values one obtains by plotting percent residual activity against [I], are independent of substrate concentration. From Henderson's equations (P.J.F. Henderson, Biochem. J. 127, 321, (1972) it follows that $IC_{50} = 1/2\ [E] + K_i$. Therefore, $IC_{50}$ values were obtained in this manner at 5 (or more) different subtilisin A concentrations. Simple graphical analyses of the data were performed: the $IC_{50}$ values were plotted against enzyme concentration, with the resultant y-intercept equal to $K_i$. $K_i$ values obtained in this manner for all inhibitors showed good agreement with those shown in Table I.

# CLAIMS

We claim:

1. A modified subtilisin inhibitor comprising a modified human standard mechanism inhibitor, wherein at least one amino acid of the reactive site of said human standard mechanism inhibitor is substituted with a different amino acid, such that the $K_I$ of said modified subtilisin inhibitor against at least one subtilisin is decreased by at least a factor of 100 relative to the unmodified human standard mechanism inhibitor.

2. A modified subtilisin inhibitor according to claim 1 wherein at least five amino acids of the reactive site of said human standard mechanism inhibitor are substituted with different amino acids.

3. A modified subtilisin inhibitor according to claim 1 wherein the reactive site of said human standard mechanism inhibitor is replaced by the reactive site of a known subtilisin inhibitor.

4. A modified subtilisin inhibitor according to claim 1 wherein the P10 to P10′ residues of the reactive site of said human standard mechanism inhibitor are replaced by the P10 to P10′ residues of the subtilisin reactive site of a known subtilisin inhibitor.

5. A modified subtilisin inhibitor according to claim 1 wherein the P6 to P5′ residues of the reactive site of said human standard mechanism inhibitor are replaced by the P6 to P5′ residues of the subtilisin reactive site of a known subtilisin inhibitor.

6. A modified subtilisin inhibitor according to claim 1, 4 or 5 wherein said human standard mechanism inhibitor comprises human pancreatic secretory trypsin inhibitor.

7. A modified subtilisin inhibitor according to claim 1, 4 or 5 wherein said human standard mechanism inhibitor comprises human mucous protease inhibitor.

8. A modified subtilisin inhibitor according to claim 4 or 5 wherein said known subtilisin inhibitor is an avian ovomucoid third domain protein.

9. A modified subtilisin inhibitor according to claim 4 or 5 wherein said known subtilisin inhibitor is a *Streptomyces* subtilisin inhibitor.

10. A modified subtilisin inhibitor according to claim 9 wherein said avian ovomucoid third domain protein is turkey ovomucoid third domain protein.

11. A method of inhibiting a target subtilisin comprising reversibly binding to a target subtilisin a modified subtilisin inhibitor comprising a human standard mechanism inhibitor wherein the reactive site of said human standard mechanism inhibitor is replaced by the reactive site of a known subtilisin inhibitor and wherein the $K_i$ of said modified subtilisin inhibitor against at least one subtilisin is decreased by at least a factor of 100 relative to the unmodified human standard mechanism inhibitor.

12. A target subtilisin inhibited by a modified subtilisin inhibitor comprising a human standard mechanism inhibitor wherein the reactive site of said human standard mechanism inhibitor is replaced by the reactive site of a known subtilisin inhibitor and wherein the $K_i$ of said modified subtilisin inhibitor against at least one subtilisin is decreased by at least a factor of 100 relative to the unmodified human standard mechanism inhibitor.

13. A nucleic acid encoding a modified human standard mechanism inhibitor, wherein at least one amino acid of the reactive site of said human standard mechanism inhibitor is substituted with a different amino acid, such that the $K_i$ of said modified subtilisin inhibitor against at least one subtilisin is decreased by at least a factor of 100 relative to the unmodified human standard mechanism inhibitor.

14. A nucleic acid according to claim 13 wherein the reactive site of said human standard mechanism inhibitor is replaced by the reactive site of a known subtilisin inhibitor.

15. A nucleic acid according to claim 14 wherein the P6 to P5' residues of the reactive site of said human standard mechanism inhibitor are replaced by the P6 to P5' residues of the subtilisin reactivie site of a known subtilisin inhibitor.

16. A nucleic acid according to claim 13 wherein said human standard mechanism inhibitor comprises human pancreatic secretory trypsin inhibitor.

17. A nucleic acid according to claim 13 wherein said known subtilisin inhibitor is an avian ovomucoid third domain protein.

18. A nucleic acid according to claim 13 wherein said known subtilisin inhibitor is a *Streptomyces* subtilisin inhibitor.

19. An expression vector comprising the nucleic acid of claim 13.

20. A host cell comprising the nucleic acid of claim 13.

21. A composition comprising:

a) a modified subtilisin inhibitor comprising a modified human standard mechanism inhibitor, wherein at least one amino acid of the reactive site of said human standard mechanism inhibitor is substituted with a different amino acid, such that the $K_i$ of said modified subtilisin inhibitor against a target subtilisin is decreased by at least a factor of 100 relative to the unmodified human standard mechanism inhibitor; and

b) a detergent containing a target subtilisin.

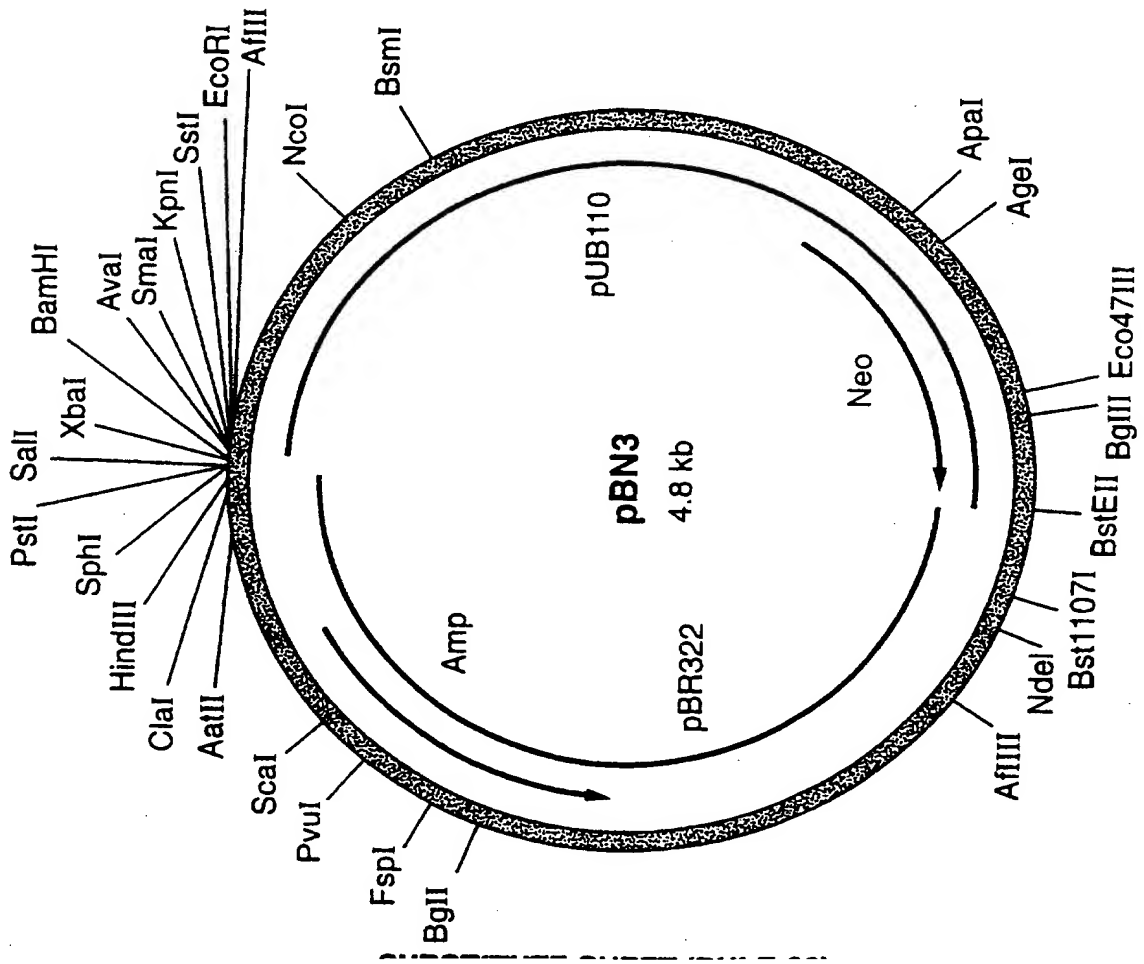
pBN3
4.8 kb
pUB110
pBR322
Amp
Neo
EcoRI
AflIII
NcoI
BsmI
ApaI
AgeI
Eco47III
BglII
BstEII
Bst1107I
NdeI
AflIII
ScaI
PvuI
FspI
BglII
AatII
ClaI
HindIII
SphI
PstI
SalI
XbaI
BamHI
AvaI
SmaI
KpnI
SstI

FIG._1

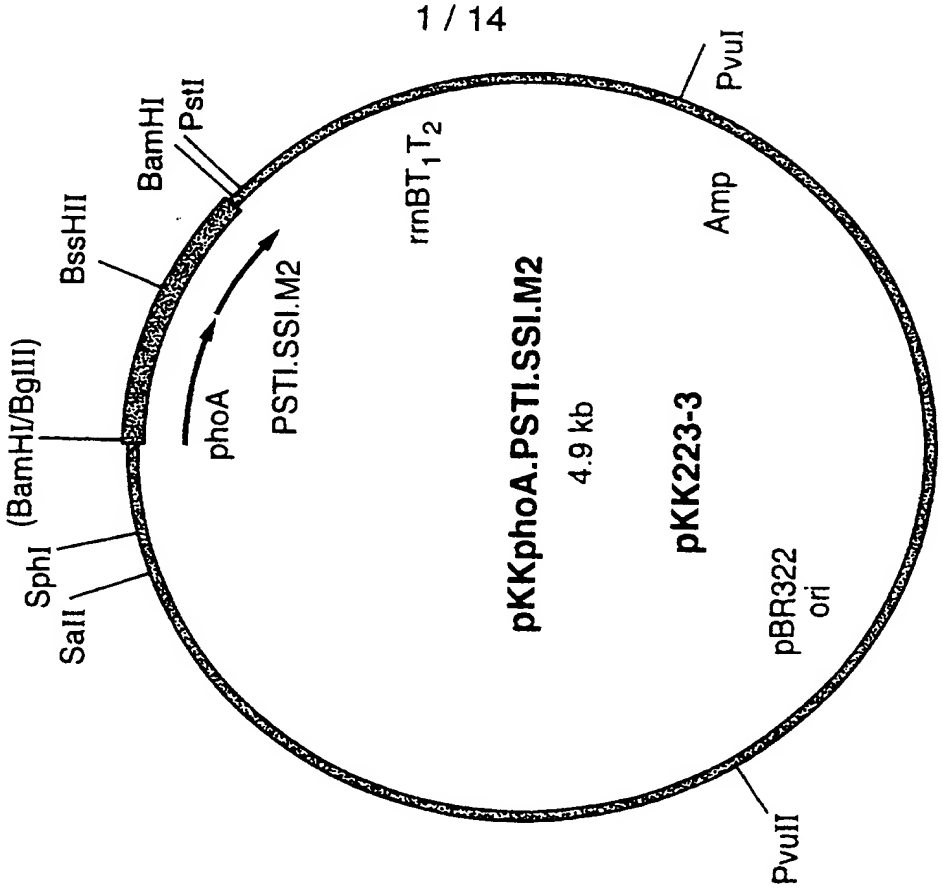
pKKphoA.PSTI.SSI.M2
4.9 kb
pKK223-3
phoA
PSTI.SSI.M2
rrnBT1T2
Amp
pBR322
ori
PvuI
PvuII
BamHI
PstI
BssHII
(BamHI/BglII)
SphI
SalI

FIG._2

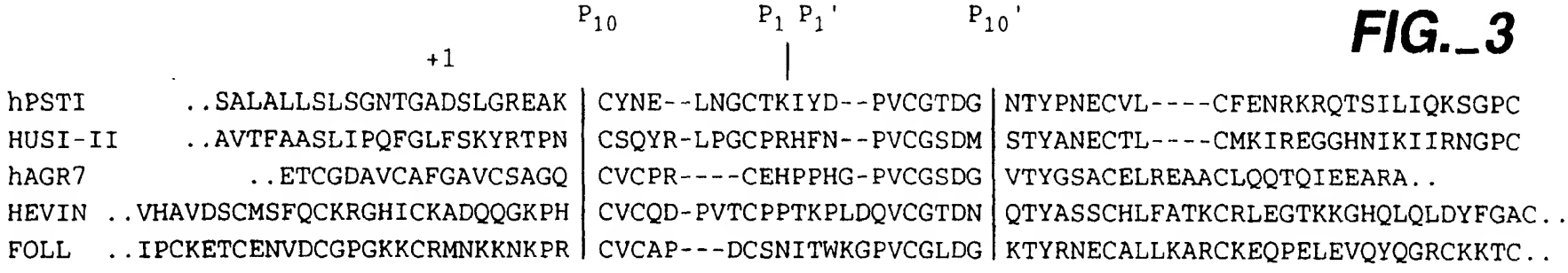


FIG._3

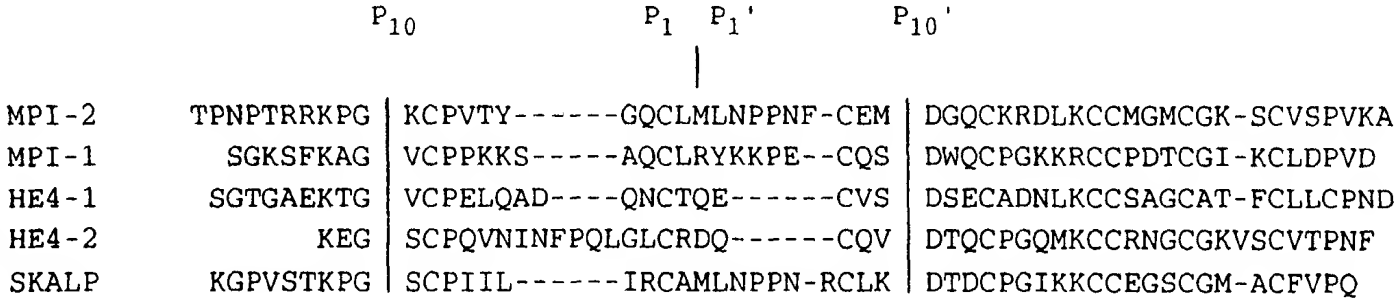


FIG._4

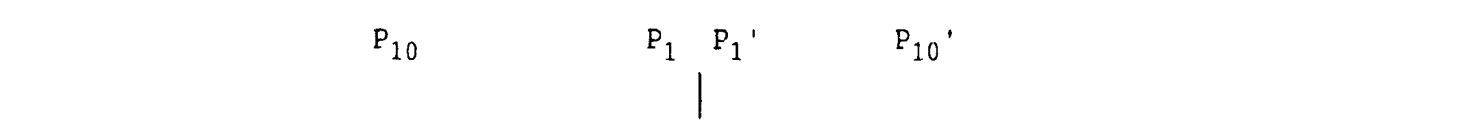


FIG._7A

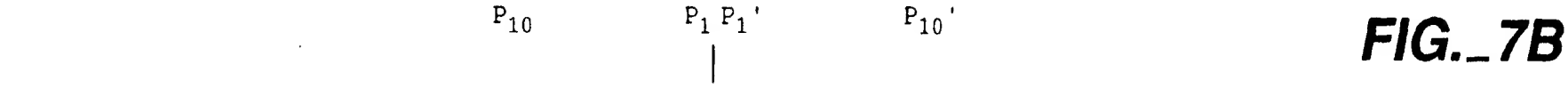


FIG._7B

| | $P_{10}$ | $P_1$ $P_1'$ | $P_{10}'$ |
|---|---|---|---|
| M. gallopavo | LAAVSVDC | SEYPKPACTLEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| C. chavaria | VATVDC | SDYPKPACTLEYMPLCGSDN | QTYSNKCSFCNAVVDSNGTLTLSHFGKC |
| A. galericulata | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| S. melanotos | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| P. strigoides | IATVDC | SDYPQPVCMGDYMPLCGSDS | KTYSNKCNFCNAVVDSNGTLTLNHFGKC |
| L. marinus | IATVDC | SDYPKPACSLDYMPLCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFEKC |
| P. plancus | VATVDC | SEYPKPVCSLEYMPLCGSDS | QTYSNECNFCNAVVDSNGTLTLSHFGKC |
| H. albicilla | VATVDC | SEYPKPVCSLEYMPLCGSDS | QTYSNECNFCNAVVDSNGTLTLSHFGKC |
| P. superciliaris | FAPVNVDC | SDHPKPACLQEQKPLCGSDN | KTYDNKCSFCNAVVDSNGTLTLSHFGKC |
| G. pucherani | LAAVDC | SEYPKPACTMEHRPLCGSDN | QTYDNKCNFCNAVVESNGTLTLSHFGKC |
| A. congensis | LAAVSVDC | SEYPKPACTMEQRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| B. thoracica | LAAVSVDC | SEYPKPECTAEERPICGSDN | KTYGNKCNFCNAVVESNGTLTLRNFGKC |
| C. mantchuricum | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| G. varius | LAAVSVDC | SEYPKPDCAAEDRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| L. bulweri | LAAVSVDC | SEYPKPTCTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| L. leucomelana | LAAVSVDC | SEYPKPACTIEYRPLCGSDN | KTYDNKCNFCNAVVESNGTLTLSHFGKC |
| L. swinhoei | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| P. muticus | AVSVDC | SEYPKPACTLEHRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| S. humiae | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| S. mikado | AVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSRFGKC |
| S. soemmerringii | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| G. chloropus | TATVDC | SGYPQPACSLEYVPFCGSDN | KTYSNKCDFCNAVADSNGTLTLSHFGKC |
| G. australis | TATVDC | SGYPQSACPQDYVPFCGSDN | KTYSNKCNFCNAVADSNGTLTLSHFGKC |
| P. albiventer | IATVDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFGKC |
| P. adelie | IATVDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFGKC |
| C. coturnix japonica | LAAVSVDC | SEYPKPACPKDYRPVCGSDN | KTYSNKCNFCNAVVESNGTLTLNHFGKC |

FIG._5A

| | $P_{10}$ | $P_1$ $P_1'$ | $P_{10}'$ |
|---|---|---|---|
| L. nycthemera | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| T. merula | IVTVDC | SDYPKPVCTLDYMPLCGSDN | TTYSNKCNFCNAVVDSNGTITLSHFGKC |
| T. blythii | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| T. caboti | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| S. novaehollandiae | TVDC | SDYPKPVCTLEEMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| N. perdicaria | VTVDC | SGYPKPACTLEYFPLCGSDN | QTYANKCTFCNAVVEKN--VTLNHLGEC |
| R. roulroul | LAAVNVDC | SGYPKPACTMESRPICGSDN | TTYSNKCNFCNAVVESNGTLTLSYPGVC |
| B. pavonina | TAMVDC | SDYPKPACTLEYMPFCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFGKC |
| C. bennetti | FATVDC | SEYPKPVCSPEYMPLCGSDS | KTYNNKCDFCSAVVESNGTLTLGHFGKC |
| A. aegyptiaca | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| A. canagicus | VATVDC | SDYPKPACTVEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| T. major | VTVDC | SGYPKPACTAEYFPLCGSDN | QTYSNKCAFCNAAVEKN--VTLGHFGEC |
| G. guira | ITTVDC | SDYPKPVCSLEYMPLCGSDN | KTYGNKCNFCNAVADSNGTLTLSHFGKC |
| O. hoazin | TATVDC | SDYPKPACLLEYMPLCGSDN | KTYDNKCSFCNAVVDSNGTLSLSHFGKC |
| N. monachus | IAIVDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCDFCNAVVDSNGTLTLSHFGKC |
| C. mexicanus | IVAVDC | SDYPKPVCSLDYMPLCGSDN | TTYNNKCIFCNAVVDSNGTITLSHFGKC |
| C. monedula | IVTVDC | SDYPKPVCTLDYMPLCGSDN | KTYSNKCNFCNAVVDSNGTITLSHFGRC |
| G. japonensis | TATVDC | SDYPKPACTLEYMPFCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFGKC |
| P. emphanum | VDC | SEYPQPACTTERRPVCGSNN | KTYSNKCNFCNAVVKSNGTLTVSHFGKC |
| C. albus | IVTVDC | SDYPRPVCTLDYMPLCGSDN | KTYSNKCNFCNAVVDSNGTITLSHFGRC |
| N. rufina | VATVNC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| R. rufescens | VTVDC | SGYPKPDCTLESFPLCGSDN | QTYSNKCAFCNAAVERN--VTLRHLGEC |
| A. americana | VATVNC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| L. diardi | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KSYDNKCNFCNAVVESNGTLTLSHFGKC |
| C. torquata | VATVDC | SDYPKPACTLEYMPLCGSDN | QTYSNKCSFCNAVVDSNGTLTLSHFGKC |
| A. arfakianus | VVIVDC | SDYPTHGCTLELKPICGSDN | QTYSNKCGFCNAVAQSNGTLTLSHFGKC |

FIG._5B

| | $P_{10}$ | $P_1$ $P_1'$ | $P_{10}'$ |
|---|---|---|---|
| C. crossoptilon | LAAVSVDC | SEYPKPACTLEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| L. leucurus | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| O. picta | FAAVSVDC | SEYPKPACTLEYRPLCGSDS | KTYANKCNFCNAVVESNGTLTLSHFGKC |
| L. gambelii | FAAVSVDC | SEYPKPACTLEYRPLCGSDN | KTYANKCNFCNAVVESNGTLTLSHFGKC |
| A. chukar | LAAVDC | SEYPKPACTLEYRPLCGSDS | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| P. colchicus | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLNRFGQC |
| S. reevesi | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSRFGEC |
| P. cristatus | LAAVSVDC | SEYPKPACTLEHRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| S. camelus | FATVDC | SDYPKPVCPLDYMPLCGSDS | KTYSNKCNFCNAVVESSGTLTLRHFGKC |
| R. americana | FATVDC | SDHPKPVCSLEYMPLCGSDS | KTYSNKCDFCNAVVESNGTLTLSHFGKC |
| P. pennata | FATVDC | SDHPKPVCSLEYMPLCGSDS | KTYSNKCDFCNAVVESNGTLTLSHFGKC |
| C. casuarius | FATVDC | SEYPKPVCSPEYMPLCGSDS | KTYNNKCDFCSAVVESNGTLTLGHFGKC |
| N. cinerascens | VTVDC | SGYPKPACTLEYFPLCGSDN | QTYANKCAFCNAVVEKN--VTLRHLGKC |
| S. humboldti | IATIDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCNFCNAVVDSNGTLILSHFGKC |
| P. sulcirostris | IATVDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCNFCNAVADSNGTLTLSHFGKC |
| A. novaehollandiae | IASVDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFGKC |
| N. nycticorax | IVTVDC | SDYPKPVCSPENMPVCGSDS | KTYSNKCDFCNAVADSNGTLTLSHFGKC |
| D. arborea | VATVDC | SDYPKPACTLEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| D. arcuata | VATVDC | SDYPKPACTLEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| D. autumnalis | VATVDC | SDYPKPACTLEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| D. bicolor | VATVDC | SDYPKPACTLEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| D. eytoni | VATVDC | SDYPRPDCTLEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| D. viduata | VATVDC | SDYPKPACTLEYMPLCGSDN | KTYGNRCNFCNAVVDSNGTLTLSHFGKC |
| C. coscoroba | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| C. atratus | VATVDC | SDYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| C. olor | VATVDC | SDYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |

FIG._5C

| | $P_{10}$ | $P_1$ $P_1'$ | $P_{10}'$ |
|---|---|---|---|
| A. anser | VATVDC | SDYPKPACTVEYMPLCGSDN | KTYDNKCNFCNAVVDSNGTLTLSHFGKC |
| A. indicus | VATVDC | SDYPKPACTVEYMPLCGSDN | KTYDNKCNFCNAVVDSNGTLTLSHFGKC |
| B. canadensis | VATVDC | SDYPKPACTVEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| C. novaehollandiae | VATVDC | SDYPKPACRMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| C. picta | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| C. cyanoptera | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| N. jubata | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| L. specularioides | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| A. capensis | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| M. cucullatus | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| O. jamaicensis | VATVDC | SDYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| O. vittata | VATVDC | SDYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGKC |
| H. indus | ITIVDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCDFCNAVVDSNGTLNLSHFGKC |
| G. coprotheres | IAIVDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCDFCNAVVDSNGTLTLSHFGKC |
| P. radiatus | IAIVDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCDFCNAVVDSNGTLTLSHFGKC |
| C. aeruginosus | IAIVDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCDFCNAFVDSNGTLSLSHFGKC |
| A. audax | IAIVDC | SDYPKPVCSLEYMPLCGSDS | KTYSNKCDFCNAVVDSNGTLTLSHFGKC |
| M. freycinet | LVTVDC | SDYPKPSCTLEDKPLCGSDN | QTYSNKCSFCNAVVDSNGTLTLSHFGKC |
| L. ocellata | VVTVDC | SGYPTHACTLELKPLCGSDN | QTYSNKCGFCNAVAQSNGTLTLSHFGKC |
| P. jacquacu | FAPVNVDC | SDHPKPACLQEQKPLCGSDN | KTYDNKCSFCNAVVDSNGTLTLSHFGKC |
| A. pipile | LAPVNVDC | SDHPKPACLQEQKPLCGSDN | KTYDNKCSFCNAVVDSNGTLTLSHFGKC |
| B. umbellus | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLNLSHFGTC |
| T. cupido | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| C. urophasianus | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| O. pictux | FAAVSVDC | SEYPKPACTLEYRPLCGSDS | KTYANKCNFCNAVVESNGTLTLSHFGKC |
| C. squamata | FAAVSVDC | SEYPKPDCTLEYRPLCGSDN | KTYANKCNFCNAVVESNGTLTLSHFGKC |

FIG._5D

| | $P_{10}$ | $P_1$ $P_1'$ | $P_{10}'$ |
|---|---|---|---|
| C. squamata | FAAVSVDC | SEYPKPDCTLEYRPLCGSDN | KTYASKCNFCNAVVESNGTLTLSHFGKC |
| C. squamata pallida | FAAVSVDC | SEYPKPDCTLEYRPLCGSDN | KTYASKCNFCNAVVESNGTLTLSHFGKC |
| C. virginianus | FAAVSVDC | SEYPKPACMSEYRPLCGSDN | KTYVNKCNFCNAVVESNGTLTLSHFGKC |
| C. montezumae | FAAVSVDC | SEYPKPACTLEYVPICGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| A. rufa | LAAVDC | SEYPKPACTLEYRPLCGSDS | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| F. afer | VSVDC | SEYPKPACTMEYRPVCGSDN | ITYGNKCNFCNAVVKSNGTLTLSHFGKC |
| F. coqui | VSVDC | SEYPKPGCTMEYRPVCGSDN | ITYGNKCNFCNAVVKSNGTLTLSHFGKC |
| F. erckelii | PAAVSVDC | SEYPKPACTMEYRPVCGSDN | ITYGNKCNFCNAVVKSNGTLTLSHFGKC |
| F. francolinus | VSVDC | SGYPKPECTAEETPLCGSDN | KTYRNKCNFCNAVVESNGTLTLRHFGKC |
| F. pondicerianus | LPAVSVDC | SEYPKPDCTTEERPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| P. perdix | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCDFCNAVVESNGTLTLNHFGKC |
| C. delegorguei | SVDC | SEYPKPACPKDYRPVCGSDN | KTYGNKCNFCNAVVESNGTLTLNRFGKC |
| C. coturnix japonica | LAAVSVDC | SEYPKPACPKDYRPVCGSDN | KTYSNKCNFCNAVVESNGTLTLNHFGKC |
| A. torqueola | VSVDC | SDYPMPACTMEYRPVCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| T. satyra | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| T. temminckii | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| L. impejanus | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| C. auritum | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| L. edwardsi | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| L. ignita | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KSYDNKCNFCNAVVESNGTLTLSHFGKC |
| P. macrolopha | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| C. wallichii | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| P. versicolor | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLNRFGQC |
| S. ellioti | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| C. amherstiae | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLNHFGKC |
| C. picta | LAAVSVDC | SEYPKPACTMEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLNHFGKC |

FIG._5E

| | $P_{10}$ | $P_1$ $P_1'$ | $P_{10}'$ |
|---|---|---|---|
| P. bicalcaratum | VSVDC | SEYPQPTCTTEHRPVCGSNN | ETYGNKCNFCNAVVKSNGTLTVSHFGKC |
| A. argus | LAAVDC | SEHPKPACTLEDRPLCGSDN | KIYSNKCDFCNAVLESNGTLTLSHFGKC |
| N. meleagris | LAAVDC | SEYPKPACTMEYRPLCGSDN | KTYDNKCNFCNAVVESNGTLTLSHFGKC |
| A. vulturinum | LAAVDC | SEYPKPACTMEYRPLCGSDN | KTYDNKCNFCNAVVESNGTLTLSHFGKC |
| A. ocellata | LAAVSVDC | SEYPKPACTLEYRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| T. sylvatica | VSVDC | SEYPKPGCMMERLPLCGSDN | KTYNDKCNFCNAVVESNGTLTLNHFGEC |
| G. carunculatus | TATVDC | SDYPKPACTLEYMPFCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFGKC |
| G. vipio | TATVDC | SDYPKPACTLEYMPFCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFGKC |
| A. virgo | TATVDC | SDYPKPACTLEYMPFCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFGKC |
| F. atra | TATVDC | SGYPQPACSLEYVPFCGSDN | KTYSNKCDFCNAVADSNGTLTLSHFGKC |
| V. spinosus | IATVDC | SDYPKPACSLDYMPLCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFEKC |
| L. ridibundus | IATVDC | SDYPKPACSLDYMPLCGSDS | KTYSNKCNFCNAVVDSNGTLTLSHFEKC |
| G. californianus | IASVNC | SDYPKPVCSLLYMPLCGSDN | KTYGNKCNFCNAVADSNGTLTLSHFGKC |
| C. renauldi | IATVDC | SDYPKPVCTLEDMPLCGSDN | ITYHNKCYFCNAVAHSNGTLTFSHFGKC |
| D. novaeguineae | IATQVDC | SEHPKPACTLDYRPICGSDS | KTYSNKCDFCNAVMDSNGTLTLSHFGKC |
| D. novaehollandiae | FATVDC | SDYPKPVCSLEYMPLCGSDS | KTYNNKCDFCNAVVESNGTLTLGHFGKC |
| E. elegans | VSIDC | SGYPKPACTLEFFPLCGSDN | QTYSNKCAFCNAAVEKNV--TLNHIGEC |
| A. semipalmeta | VATVDC | SEYPKPACTLEYMPLCGSDN | QTYSNKCNFCNAVVDSNGTLTLSHFGKC |
| T. radjah | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| A. platyrhynchos | VATVDC | SGYPKPACTMEYMPLCGSDN | KTYGNKCNFCNAVVDSNGTLTLSHFGEC |
| O. vetual | FAPVNVDC | SDHPKPACLQEQKPICGSDN | KTYDNKCSFCNAVVDSNGTLTLSHFGKC |
| C. squamata castanogastric | FAAVSVDC | SEYPKPDCTLEYRPLCGSDN | KTYANKCNFCNAVVESNGTLTLSHFGKC |
| L. californicus | FAAVSVDC | SEYPKPACTLEYRPLCGSDN | KTYASKCNFCNAVVESNGTLTLSHFGKC |
| G. lafayettii | LAAVSVDC | SEYPKPDCTAEDRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |
| G. sonneratii | LAAVSVDC | SEYPKPDCTVEDRPLCGSDN | KTYGNKCNFCNAVVESNGTLTLSHFGKC |

FIG._5F

| | $P_{10}$ | $P_1$ $P_1'$ | $P_{10}'$ |
|---|---|---|---|
| SSI | LTCAPGPSGTHPAAGSACADLAAVGGDLNAL | TRG---EDVMCPMVYDPVLLTVD | GVWQGKRVSYERVFSNECEMNAHGSSVFAF |
| SIL1 | LRCLPVG-GDHPAPEKACAALREAGGDPAAL | PRYVEDTGRVCTREYRPVTVSVQ | GVWDGRRIDHAQTFSNSCELEKQTASVYAF |
| SIL2 | LTCAPKATGTHPAADAACAELRRAGGDFDAL | SAA---DGVMCTREYAPVVVTVD | GVWQGRRLSYERTFANECVKNAGSASVFTF |
| SIL3 | LTCAPTASGTHPAADAACAELRAAHGDPSAL | AAD---DAVMCTREYAPVVVTVD | GVWQGRRLSYERTFANECVKNAGSASVFTF |
| SIL4 | LTCAPTSSGTHPAASAACAELRGVGGDFAAL | KAR---DDVWCNKLYDPVVVTAQ | GVWQGQRVSYERTFGNSCERDAVGGSLFAF |
| STI-1 | LTCAPTASGTHPAAAAACAELRAAHGDPSAL | AAE---DSVMCTREYAPVVVTVD | GVWQGRRLSYERTFANECVKNAGSASVFTF |
| STI-2 | LNCAPTASGTHPAPALACADLRGVGGDIDAL | KAR---DGVICNKLYDPVVVTVD | GVWQGKRVSYERTFGNECVKNSYGTSLFAF |
| PSN | LNCAPTASGTHPAALQACAELRGAGGDFDAL | TVR---GDVACTKQFDPVVVTVD | GVWQGKRVSYERTFANECVKNSYGMTVFTF |
| API-2c' | LTCAPGPSGTHPAADSACADLAAVGGDLDAL | TRS---EGVMCPMIYDPVLLTVD | GVWQGERVSYERVFSNECEMNAHGSSVLAF |
| SIL8 | VSCAPGARGTHPDPKAACAALKSTGGAFDRL | LSE-PNPDRACPMHYAPVTVSAV | GVWEGRRVAWDHTFANSCTMAATLDGNAVF |
| SIL10 | LTCAPKAAGTHPAAGAACAELRGVGGDFDAL | TAR---DGVMCTKQYDPVVVTVE | GVWQGKRVSYERTFSNDCMKNAYGTGVFSF |
| SIL13 | LNCAPSASGTHPAPALACAELRAAGGDLDAL | AGP---ADTVCTKQYAPVVITVD | GVWQGKRVSYERTFANECVKNASGSSVFAF |
| SIL14 | LTCAPRPSGTHPVAGSACAELRGVGGDVHAL | TAT---DGVMCTKQYDPVVVTVD | GVWQGRRVSYERTFSNECVKNAYGSGVFAF |
| STVCI | LTCMPKADGTHPNTRGACAQLRLAGGDFEKV | TKI--KEGTACTREWNPSVVTAE | GVWEGRRVSFERTFANPCELKAGKGTVFEF |

FIG._6

*Bam*HI

GCGGCGGATCCCGGCAACACTGGAGCTGACTCCTTGGGAAGAGAGGCCAAATGTTACAAA

A A D P G N T G A D S L G R E A K C Y K>

GAACTTAATGGATGCACCAAGATATATGACCCCGTCTGTGGTACCGATGGAAATACTTAT

E L N G C T K I Y D P V C G T D G N T Y>

CCCAATGAATGCGTGTTATGTTTTGAAGGTCGGAAACGCCAGACTTCTATCCTCATTCAA

P N E C V L C F E G R K R Q T S I L I Q>

*Eco*RI

AAATCTGGGCCTTGCGCCCCCGACTACAAGGACGACGACGACAAGTGAGGAATTCCGG

K S G P C *A P D Y K D D D D K* *

**FIG._8**

$P_1 P_1'$

|

LQCPKNSEVRNSPCPRTCNDPYGQNSCITVIRETCHCKGELVFDSDSICVPISQC

**FIG._13**

*Bam*HI

GCGGCGGATCCCTCCCCATGCCATGGAGTGCAATGTGCATTTGGGGCAGTATGCACAGTG

A A D P S P C H G V Q C A F G A V C T V

AAGAACGGGAAAGCCGAGTGCGAGTGCCAGCGGGTGTGCTCCGGCATCTACGATCCTGTG

K N G K A E C E C Q R V C S G I Y D P V

TGCGGCAGTGACGGTGTCACTTACGGCAGTGTGTGCGAGCTGGAATCCATGGCCTGTACC

C G S D G V T Y G S V C E L E S M A C T

CTTGGGCGGGAAATCCAAGTGGCCCGCAGAGGACCCTGTGCCCCCGACTACAAGGACGAC

L G R E I Q V A R R G P C *A P D Y K D D*

*Eco*RI

GACGACAAGTGAGGAATTCCGGAGCC

*D D K* *

**FIG._14**

(BamHI/BglII)
GGATCTTTTGGAGATTATCGTCACTGCAATGCTTCGCAATATGGCGCAAAATGACCAACA

GCGGTTGATTGATCAGGTAGAGGGGGCGCTGTACGAGGTAAAGCCCGATGCCAGCATTCC

TGACGACGATACGGAGCTGCTGCGCGATTACGTAAAGAAGTTATTGAAGCATCCTCGTCA

GTAAAAAGTTAATCTTTTCAACAGCTGTCATAAAGTTGTCACGGCCGAGACTTATAGTCG

CTTTGTTTTTATTTTTTAATGTATTTGTACATGGAGAAAATAAAGTGAAACAAAGCACTA
M K Q S T

*Bss*HII
TTGCACTGGCACTCTTACCGTTACTGTTTACCCCTAGCGCGCAGGCGGCAGGCAACACTG
I A L A L L P L L F T P S A Q A A G N T

*Avr*II
GAGCTGACAGCCTAGGAAGAGAGGCCAAATGTTACAATGAAGACGTAATGTGCCCTATGG
G A D S L G R E A K C Y N E **D V M C P M**

*Kpn*I
TGTATGATCCTGTCTGTGGTACCGATGGAAATACTTATCCCAATGAATGCGTGTTATGTT
**V Y D P V** C G T D G N T Y P N E C V L C

TTGAAGGTCGGAAACGCCAGACTTCTATCCTCATTCAAAAATCTGGGCCTTGCGGAGGTG
F E G R K R Q T S I L I Q K S G P C *G G*

*Bam*HI *Pst*I
GCGGGAGTGATTACAAGGACGATGACGATAAATGATAAGGATCCTGCAG
*G G S D Y K D D D D K* *

**FIG._9A**

GNTGADSLGREAKCYNE**DVMCPMVYDPV**CGTDGNTYPNECVLCFEGRKRQTSILIQKSGPC

**FIG._9B**

*Bss*HII
GCGCGCAGGCGGCAGGCAACACTG
A O A A G N T

*Avr*II *Sph*I
GAGCTGACAGCCTAGGAAGAGAGGCCAAATGTTACAATGAAAAGCCTGCATGCACCCTGG
G A D S L G R E A K C Y N E **K P A C T L**

*Kpn*I
AGTATCGTCCTGTCTGTGGTACCGATGGAAATACTTATCCCAATGAATGCGTGTTATGTT
**E Y R P** V C G T D G N T Y P N E C V L C

TTGAAGGTCGGAAACGCCAGACTTCTATCCTCATTCAAAAATCTGGGCCTTGCGGAGGTG
F E G R K R Q T S I L I Q K S G P C *G G*

*Bam*HI *Pst*I
GCGGGAGTGATTACAAGGACGATGACGATAAATGATAAGGATCCTGCAG
*G G S D Y K D D D D K* *

**FIG._10A**

GNTGADSLGREAKCYNE**KPACTLEYRPL**CGTDGNTYPNECVLCFEGRKRQTSILIQKSGPC

**FIG._10B**

BssHII
GCGCGCAGGCGGCAGGCAACACTG
A O A A G N T

*Avr*II
GAGCTGACAGCCTAGGAAGAGAGGCCAAATGTTACAATGAAGCCGCGGCCTGCCCTCTGG
G A D S L G R E A K C Y N E **A A A C P L**

*Kpn*I
TGTATCGTCCTGTCTGTGGTACCGATGGAAATACTTATCCCAATGAATGCGTGTTATGTT
**V Y R P V** C G T D G N T Y P N E C V L C

TTGAAGGTCGGAAACGCCAGACTTCTATCCTCATTCAAAAATCTGGGCCTTGCGGAGGTG
F E G R K R Q T S I L I Q K S G P C *G G*

*Bam*HI *Pst*I
GCGGGAGTGATTACAAGGACGATGACGATAAATGATAAGGATCCTGCAG
*G G S D Y K D D D D K* *

**FIG._11A**

GNTGADSLGREAKCYNE**AAACPLVYRPV**CGTDGNTYPNECVLCFEGRKRQTSILIQKSGPC

**FIG._11B**

*Eco*RI *B. subtilis aprE* promoter --->

GAATTCCTCCATTTTCTTCTGCTATCAAAATAACAGACTCGTGATTTTCCAAACGAGCTT
TCAAAAAAGCCTCTGCCCCTTGCAAATCGGATGCCTGTCTATAAAATTCCCGATATTGG
TTAAACAGCGGCGCAATGGCGGCCGCATCTGATGTCTTTGCTTGGCGAATGTTCATCTTA
TTTCTTCCTCCCTCTCAATAATTTTTTCATTCTATCCCTTTTCTGTAAAGTTTATTTTTC
AGAATACTTTTATCATCATGCTTTGAAAAAATATCACGATAATATCCATTGTTCTCACGG
AAGCACACGCAGGTCATTTGAACGAATTTTTTCGACAGGAATTTGCCGGGACTCAGGAGC
ATTTAACCTAAAAAAGCATGACATTTCAGCATAATGAACATTTACTCATGTCTATTTTCG
TTCTTTTCTGTATGAAAATAGTTATTTCGAGTCTCTACGGAAATAGCGAGAGATGATATA
CCTAAATAGAGATAAAATCATCTCAAAAAAATGGGTCTACTAAAATATTATTCCATCTAT
TACAATAAATTCACAGAATAGTCTTTTAAGTAAGTCTACTCTGAATTTTTTTAAAAGGAG

*B. subtilis* <---|---> *B. amyloliquefaciens*

AGGGTAAAGAGTGAGAAGCAAAAAATTGTGGATCAGTTTGCTGTTTGCTTTAGCGTTAAT
M R S K K L W I S L L F A L A L I

*Bss*HII

CTTTACGATGGCGTTCGGCAGCACATCCAGCGCGCAGGCGGCAGGGAAATCAAACGGGGA
F T M A F G S T S S A Q A A G K S N G E

AAAGAAATATATTGTCGGGTTTAAACAGACAATGAGCACGATGAGCGCCGCTAAGAAGAA
K K Y I V G F K Q T M S T M S A A K K K

AGATGTCATTTCTGAAAAAGGCGGGAAAGTGCAAAAGCAATTCAAATATGTAGACGCAGC
D V I S E K G G K V Q K Q F K Y V D A A

TTCAGCTACATTAAACGAAAAAGCTGTAAAAGAATTGAAAAAAGACCCGAGCGTCGCTTA
S A T L N E K A V K E L K K D P S V A Y

SphI PacI BamHI PstI

CGTTGAAGAAGATCACGTAGCGCATGCTTAATTAAGGATCCTGCAGCTCAGTAAAACATA
V E E D H V A.....A A Q Stop

*B. amyloliquefaciens* terminator --->

AAAAACCGGCCTTGGCCCCGCCGGTTTTTTATTATTTTTCTTCCTCCGCATGTTCAATCC
GCTCCATAATCGACGGATGGCTCCCTCTGAAAATTTTAACGAGAAACGGCCGGTTGACCC
GGCTCAGTCCCGTAACGGCCAAGTCCTGAAACGTCTCAATCGCCGCTTCCCGGTTTCCGG

HindIII

TCAGCTCAATGCCGTAACGGGCGGCCGCAAGCTT

**FIG._12**

# INTERNATIONAL SEARCH REPORT

Inter onal Application No
PC /US 96/17153

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/15 C07K19/00 C07K14/81 C12N9/54 C12N1/21
C11D3/386

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | WO 93 00418 A (GENENCOR INT) 7 January 1993 | 1-8, 10-17, 19-21 |
| X | see the whole document | 12 |
| Y | WO 92 03529 A (NOVONORDISK AS) 5 March 1992 | 1-7,9, 11-16, 18-21 |
| X | see the whole document | 12 |

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 24 February 1997 | 11. 03. 97 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. ( + 31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax ( + 31-70) 340-3016 | Van der Schaal, C |

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | DATABASE CHEMABS<br>CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US<br>WITTEMAN, AGNES M. ET AL: "Food allergens in house dust"<br>XP002025987<br>see abstract<br>& INT. ARCH. ALLERGY IMMUNOL. (1995), VOLUME DATE 1995, 107(4), 566-8 CODEN: IAAIEG;ISSN: 1018-2438,<br>1995,<br>--- | 1-21 |
| Y | DATABASE MEDLINE<br>Accession no. 88088388, 1988<br>D. BERNSTEIN ET AL: "Clinical and immuniological studies among egg-processing workers with occupational asthma"<br>XP002025988<br>see abstract<br>& JOURNAL OF ALLERGY AND CLINICAL IMMUNOLOGY,<br>vol. 80, no. 6, December 1987,<br>pages 791-797,<br>--- | 1-21 |
| Y | PROTEIN ENGINEERING,<br>vol. 8, no. 1, 1 January 1995,<br>pages 45-52, XP000500392<br>SZARDENINGS M ET AL: "HIGHLY EFFECTIVE PROTEASE INHIBITORS FROM VARIANTS OF HUMAN PANREATIC SECRETORY TRYPSIN INHIBITOR (HPSTI): AN ASSESSMENT OF 3-D STRUCTURE-BASED PROTEIN DESIGN"<br>see the whole document<br>especially page 45, right column, paragraph 2<br>--- | 1-21 |
| Y | WO 91 00912 A (MASSACHUSETTS INST TECHNOLOGY) 24 January 1991<br>see the whole document<br>especially page 17, lines 15-22<br>--- | 1-21 |
| Y | JOURNAL OF BIOCHEMISTRY,<br>vol. 109, no. 3, 1 March 1991,<br>pages 377-382, XP000145568<br>SHUICHI KOJIMA ET AL: "INHIBITION OF SUBTILISIN BPN' BY REACTION SITE P1 MUTANTS OF STREPTOMYCES SUBTILISIN INHIBITOR1"<br>see the whole document<br>--- | 9,18 |

-/--

# INTERNATIONAL SEARCH REPORT

Inter ·ional Application No
PC /US 96/17153

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 83, 1986, WASHINGTON US, pages 6692-6696, XP002019052 R. THOMPSON AND K. OHLSSON: "Isolation, properties, and complete amino acid sequence of human secretory leukocyte protease inhibitor...." see figure 7 --- | 7 |
| A | WO 95 17885 A (UNIV RUTGERS ;ANDERSON STEPHEN (US); RYAN RAYMOND (US)) 6 July 1995 see page 6-8 ----- | |

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

nformation on patent family members

Inter- ·ional Application No
PC·/US 96/17153

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 9300418 A | 07-01-93 | US 5178789 A | 12-01-93 |
| | | CA 2111970 A | 07-01-93 |
| | | EP 0591445 A | 13-04-94 |
| | | JP 7501350 T | 09-02-95 |
| WO 9203529 A | 05-03-92 | DE 69101557 D | 05-05-94 |
| | | DE 69101557 T | 14-07-94 |
| | | EP 0544777 A | 09-06-93 |
| | | ES 2062812 T | 16-12-94 |
| | | JP 6500142 T | 06-01-94 |
| WO 9100912 A | 24-01-91 | NONE | |
| WO 9517885 A | 06-07-95 | US 5550213 A | 27-08-96 |
| | | AU 1515295 A | 17-07-95 |